Classroom Acoustics - Controlling the Cafe Effect... is the Lombard Effect the key?

James Whitlock (1) and George Dodd (2)

(1) Marshall Day Acoustics, Auckland, New Zealand
(2) Acoustics Research Centre, University of Auckland, New Zealand

ABSTRACT

New Standards for classroom design indicate that classrooms for young children should be built to criteria which in some respects – e.g. reverberation time – differ markedly from the traditional criteria for adults. Research in New Zealand has investigated the reasons for these differences. To identify why the reverberation needs of children and adults for speech perception are so different we have measured speech integration times for adults and children using a novel technique of reversed-segmented speech to obviate the confounding effects of differing language abilities in children. The values found for the two groups are significantly different. Background noise is a critical feature in classrooms and, with present day teaching emphasizing interactive learning; it is the activity noise of the children which is the main component. When groups of children are working independently in the same classroom the "cafe effect" produces a rising noise level as children compete to be heard. We suggest this results from the Lombard effect. From measurements of this effect in children and adults we have been able to successfully predict the activity noise measured in a classroom, and compare differences in susceptibility and magnitude between these two groups.

INTRODUCTION

This paper is a continuation of that presented at the last NZAS conference, based on recent investigations into speech intelligibility requirements of children in a classroom environment (Whitlock, 2003). We revisit the conclusions regarding integration time of speech and augment the conclusions regarding the Lombard Effect, in light of subsequent research on adults (Francis, 2005) which further highlights the differences between children and adults.

This new work strengthens the hypothesis that children's' speech intelligibility requirements must be the controlling factor in classroom acoustical design, and that some existing design standards for reverberation time (ANSI, 2002) are not stringent enough. This standard recommends a 0.6 second reverberation time (RT), whereas the New Zealand Classroom Acoustics Research Ground (NZCRG) found that classrooms which exhibited RT's around this value were generally rated as poor, and suggests 0.4 second is a more suitable value (Wilson et al., 2002).

Reverberation time is the main acoustic parameter which impacts on the speech intelligibility and signal-to-noise ratio as it heightens the presence of late-energy in speech, which is destructive to intelligibility, and exacerbates background noise via the café effect.

In simple terms the café effect is the tendency, inside a reverberant space, for noise to 'breed' noise. Generally the noise is the conversation of individual groups of occupants, who subconsciously compete with one another for signal-to-noise ratio such that they can be heard and understood by their peers.

When taking account of the café effect, it is common to assume the phenomenon is wholly governed by ones perceived requirements for social interaction i.e. "I must speak louder so my friends can understand me." However we hypothesise a second motivation, which is that the speaker raises their voice level in order to hear themselves. This mechanism is accounted for by the Lombard Effect (Lombard, 1911), which is a well established and widely documented phenomenon, though not in relation to the café effect.

Could the Lombard Effect be the trigger for the café effect, and if so, if one could control their susceptibility to the Lombard Effect, could the café effect be eliminated?

THE PHENOMENA

INTEGRATION TIME OF SPEECH – A REVIEW

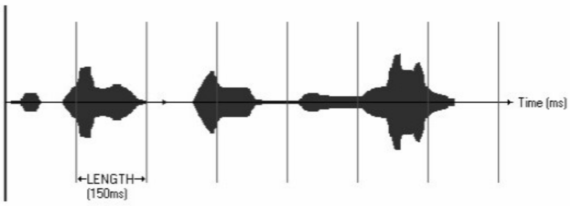
For a listener in a reverberant space, there is a point in time after the arrival of direct sound where reflections begin to become less than 100% useful for speech intelligibility. This time delay we have termed "The Integration Time of Speech". It is the threshold which divides fully useful and partially detrimental sound energy. For adults, this point is generally taken as being 50 milliseconds. (Henry, 1851; Miller, 1948; Haas, 1972; Whitlock, 2001).

We hypothesise that a reason why young children benefit from a lower RT than is appropriate for adults, is that their hearing systems are not fully mature so their ability to utilise early reflections is reduced. In particular we suggest that this might be evidenced by a shorter integration time of speech.

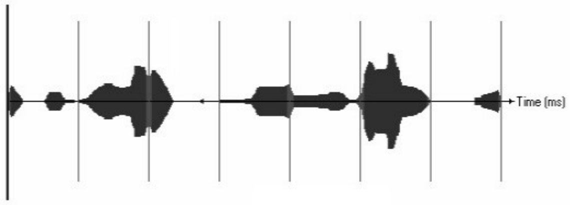
To test our hypothesis it was important that we use a speech test signal and we have devised a novel technique suggested by an effect demonstrated by Saberi and Perrott (Saberi & Perrott, 1999).

The technique is based on using reversed-segmented speech. If we view the integration phenomenon as one of a simple summing of the sound energy entering the ear over a period of time then it should be possible to "chop" a speech train into segments equal to the time of this integration and, providing the order of the segments is maintained, reverse each segment without losing the intelligibility of the speech.

This is quite a dramatic effect when experienced for the first time. Figure 1a illustrates the speech signal prior to processing, and Figure 1b illustrates the processed signal.

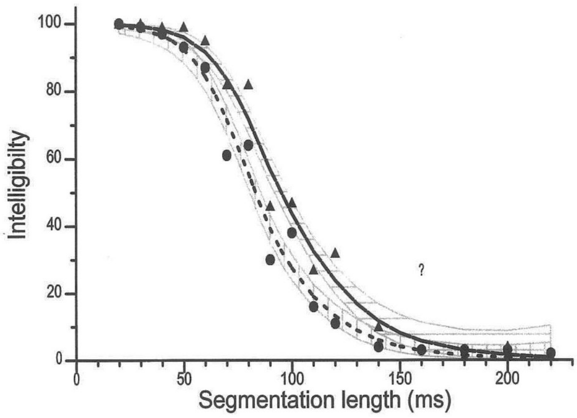


Source: (Whitlock, 2003)
Figure 1a. Unprocessed speech stream



Source: (Whitlock, 2003)
Figure 1b. Reversed segmented speech stream. Sentence chopped into segments with each segment reversed in time

A group of 15 adults and another of 18 children (7-9½ years) were individually presented in an anechoic room with BKB sentence lists (Bench, Kowal & Bamford, 1979) which had been segmented and reversed using a range of segmentation times (20-220ms). A presentation level of 67 dB(A) was chosen to ensure ease of audibility. The subjects repeated the sounds they heard and the correctly-perceived phonemes were scored by the tester. Figure 2 shows a comparison of curve-fitted results for the child and adult groups. The difference between the groups is significant at the 5% level (except for segmentation times at the extremes where no difference is to be expected).



Source: (Whitlock, 2003)
Figure 2 Intelligibility scores for the children (circles) and adults (triangles)

Taking the 98% correct score as the boundary between full and partial integration confirms our hypothesis that the integration times for speech in children and adults are markedly different (35ms and 50ms respectively). That is, children have an integration time approximately 70% of the adult value.

The Café Effect

The café effect is an extremely common, yet under-diagnosed acoustical phenomenon. Any noisy restaurant or busy café is likely to have fallen foul of its trickery, and the frustrated occupants can have practically no control whatsoever over the situation.

Possibly the most crucial arena for the café effect though is the classroom, where speech intelligibility and adequate signal-to-noise ratio are paramount to learning. As mentioned above, primary schools are particularly at risk because of the language abilities of its young pupils (and hence their need for clear speech), and because of the prevalence of group work activities.

The New Zealand Classroom Acoustics Research Group (Wilson et al, 2002) found in their teacher survey that group work is the most common method of classroom activity, accounting for 38% of teaching styles utilised. It is during these group work sessions, where students communicate with one another through so called "incidental learning" (Flexer, 1999 (cited in Wilson et al., 2002), that the café effect occurs.

The ultimate noise level is likely governed by the acoustical properties of the room; suffice to say that spaces with poor acoustic treatment (i.e. reverberative or live) exacerbate the effect and enhancing the disturbance of the speakers. This once again alludes to the vital importance of reverberation time in classrooms.

The phenomenon is sometimes referred to (particularly in the U.S.) as the Cocktail-Party Effect (MacLean, 1959), as this is obviously a social situation in which the effect is highly noticeable. However, our understanding is that the term Cocktail-Party Effect relates to a different phenomenon, as described by Cherry (Cherry, 1953) whereby an occupant in a busy room is able to selectively 'tune in' to and understand another speaker's voice over the dominant background level, even if that speaker is not in the immediate vicinity.

The Lombard Effect

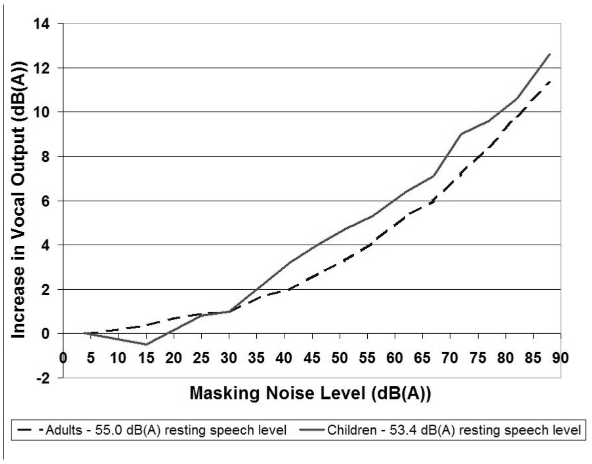
The psychoacoustical effect referred to as the Lombard Effect is so-called because of the pioneering work of Etienne Lombard (Lombard, 1911). It describes the tendency for a speaker to raise their voice in the presence of background noise. Lombard suggests it occurs so that the speaker can (a) hear themselves and (b) feel that they are communicating adequately with a listener or listeners. It is an effect which some few people can overcome to some degree by conscious control of their voice level, but the vast majority of people are unable to succeed at this (Pick et al., 1989).

We suggest that the Lombard Effect is largely responsible for the occurrence of the café effect, a view which is shared by Lubman and Sutherland (Lubman & Sutherland, 2002).

Two recent studies have investigated this effect, firstly in children (Whitlock, 2003) and then in adults (Francis, 2005). As with the Integration Time of speech experiment, we hope to expose a significant difference between children and adults, thereby vindicating the theory that classrooms must be designed with the specialist acoustic needs of children at the fore.

The same group of 18 children (7-9½ years) involved in the integration time experiment, and an ancillary group of 30 adults (20-61 years) were tested in an anechoic room.

Each subject was asked to read out loud a story from a book or magazine whilst broadband masking noise was delivered to them via insert earphones at incremented levels ranging from 4 to 88 dB(A). The reading material was chosen to be unchallenging insofar as was possible, so that the subjects would be able to read without pausing due to difficulty with words. Their resulting voice level was measured at each increment, see Figure 3 below, which shows the rise in speech level with increased masking noise, with respect to the average measured base voice level i.e. for no masking noise (53.4 dB(A) in children and 55 dB(A) in adults:



Source: (Francis, 2005)

Figure 3 Lombard Effect in Children vs Adults (with respect to base speech levels)

For both children and adults, the results of this experiment show a strong Lombard reflex and a consistent rise in speech level for masking noise above 15 dB(A) in children, and above 4 dB(A) (i.e. for all masking levels presented) in adults.

From these 'trigger' masking noise levels to the maximum 88 dB(A) level used, there was an average rise in speech level of 13.9 dB(A) in children and 11.3 dB(A) in adults. Or alternatively, a 'Lombard Coefficient' (i.e. rise in voice level per decibel of background noise level) of 0.19 dB/dB in children, and 0.13 dB/dB in adults. That is, the adults have a Lombard Effect approximately 68% of the children value.

This margin is remarkably similar to that found in the Integration Time of Speech experiment.

We note that for both groups, the slope in Figure 3 increases for masking noise levels above 65 dB(A), indicating a greater rise in speech level. Keeping activity noise levels below this point could avoid the transition into this area of heightened response to background noise.

Using the Lombard Coefficient data obtained for the child experiment we developed the following prediction model for activity noise in a classroom (Whitlock, 2003):

Let the base (resting) voice level be B dB(A)

Let the number of children speaking be N

Let the starting level for the Lombard effect be S dB(A)

Let the Lombard Coefficient be L dB/dB

Let the volume of the classroom be V, and

Let the reverberation time of the classroom be T

If the final level is F dB(A) with N children, then:

$$F = \frac{B - SL + 10\log N - R}{(1 - L)} \quad (1)$$

Where: $R = 20\log\left(0.057\sqrt{V/T}\right)$ (i.e. the rev. radius)

Assuming a classroom with a volume of 200m^3 and reverberation time of 0.6s, containing 30 students, in which activity noise is the *only* source of background noise, if we were to consider a group work activity in which the students were working in pairs, and only one of the pair was talking at any one time, there would be 15 students generating the activity noise. If each student were speaking at the average resting level of 53.4 dB(A), then (from equation 1), the generated noise level would be approx. 74 dB(A).

These values correlate remarkably with activity noise levels of between 72 and 77 dB(A) measured in other recent classroom acoustics studies (Shield & Dockrell, 2003; Whitlock, 2003; MacKenzie, 1999 (cited in Wilson et al., 2002). We feel therefore that this may be a valid model for classroom activity noise.

This correspondence suggests that the Lombard Effect may be wholly responsible for the Café Effect, however as the trigger level in both groups is so low, it seems unlikely that the Café Effect can be avoided altogether by controlling background noise.

CONCLUSIONS

We have investigated two physiological phenomena related to the auditory sensitivity of reverberation in a space, with a particular focus on primary school children in the classroom.

In both the Integration Time of Speech, and the Lombard Effect experiments, children were found to have significantly more detrimental responses to that of adults. Therefore the presence of reverberation in a space is shown to be more damaging to children in the areas of speech intelligibility and response to background noise.

In conjunction with the findings and suggested criteria in other research in this area, we can take a step closer to designing an optimum acoustic environment for primary school children, such that their speech intelligibility is maximised, which is a clear prerequisite if their full learning potential is to be realised.

FURTHER INVESTIGATIONS

Since a Lombard reflex was seen in both groups for very low masking noise levels (i.e. well below levels one would practically expect in a populated area), there is no evidence to support the notion of a practical trigger level for the Café Effect. However we suggest that further research be conducted on the Lombard effect using various types of masking noises (pure tones, speech babble), and in different test rooms. Our tests were conducted in a quiet anechoic chamber, and similar tests in more typical conditions may elicit different results.

Performing these same experiments on hearing impaired children is essential for future research, since hearing impaired students are mainstreamed in New Zealand and the needs of these students *must* be provided for. It is for this reason that the NZCRG supported the provision of FM aids for all hearing impaired children. Designing classrooms for their more critical speech intelligibility needs will improve things for normally hearing children as well.

There are additional factors relating to speech intelligibility, and the verbal communication stream in the classroom which have, through our investigations, become evident.

Voice Quality

The fundamental principle in the noise control aspects of acoustics is first and foremost addressing the noise source. This logic could be applied to didactic learning periods in the classroom, as even good acoustic design is only a conduit for the voice, and cannot enhance a poor signal.

Vocal factors such as loudness, pitch, clarity, articulation, timbre and meter could be measured and the discoveries of the relative importance of these – and possibly other factors – could be used to shape voice training or speech production techniques for teachers.

Perhaps an objective measure of voice quality could be developed to identify any desirable features in the voices of teachers who command a high degree of class control. We understand that such a measure has been investigated for the voice quality of football announcers in the UK, but we have been unable to find any details of this study.

Sound field / Teacher voice amplification systems

A developing trend for achieving higher S/N ratios is the introduction of so-called 'sound-field amplification' or "teacher voice amplification" systems into classrooms, where the teacher wears a wireless microphone and his or her voice is amplified and delivered to the class via an array of loudspeakers fixed to the walls.

Although this should increase signal-to-noise ratio appropriately, we believe it fails to address the central issue of poor room acoustics, and certainly does not improve the *listening environment* for the students.

Further, we hypothesise that if students were to experience a system like this from early on in their schooling, they may have the development of essential listening skills (such as localization & discrimination) hindered, as the amplification system removes the need for really 'attending' to the speaker. Students who change schools or classes from one fitted with a sound-field amplification system to one which was not may also experience serious disruption at critical periods of their education.

The noise levels produced by a system such as this may also significantly increase the daily noise dose of a child and potentially create intrusive noise problems for near-by classrooms. Furthermore, teachers may come to depend on such a system for communication and use it overmuch, out of convenience.

We take the view that the natural acoustics of the room should be improved as much as possible so that the environment lends itself to good speech communication without the need for aids i.e. the noise should be decreased, rather than the signal increased. This means that the resolving of background noise issues and the reinforcement of speech intelligibility for both student-teacher and student-student interaction through good acoustic design is paramount.

A study is currently underway to find out whether such systems do increase the direct-to-reverberant ratio of the teachers voice sufficiently to counter a poor reverberation time.

REFERENCES

ANSI, 2002, S12.60-2002 - Acoustical performance criteria, design requirements, and guidelines for Schools, American National Standards Institute

Bench, J., Kowal, A. & Bamford, J. 1979 *The BKB Sentence Lists,* Brit. J. of Aud. 13, pp. 18-112

Cherry, E.C. 1953, Some Experiments on the recognition of speech, with one and two ears, J. Acous. Soc. Am. 25, pp. 975-979

Francis, R. 2005 *The Influence of the Lombard Effect on Speech Level in Adults,* Research Paper, School of Music, University of Auckland

Haas, H. 1972, The Influence of a Single Echo on the Audibility of Speech, Aud. Eng. Soc. Journal 20, pp. 146-159

Henry, J. 1851, *On the Limit of Perceptibility of a Direct and Reflected Sound,* in Scientific Writings of Joseph Henry, pp. 295-296, Smithsonian Institution, Washington

Lombard, E. 1911, *Le signe de l`élévation de la voix,* Ann. Maladies Orielle, Larynx, Nez, Pharynx 37, pp. 101-119 [Translated into English by T. Scelo (2003)]

Lubman, D. & Sutherland, L. 2002, *Role of Soundscape in Children's Learning,* Proceedings of First Pan-American/Iberian Meeting on Acoustics, Cancun, Mexico

MacLean, W.R. 1959, *On the Acoustics of Cocktail Parties,* J. Acoust. Soc. Am. 31, pp. 79-80

Maddox, J. 1994, *Cocktail party effect made tolerable,* Nature Vol. 369, p. 517

Miller, G.A. 1948, *The Perception of Short Bursts of Noise,* J. Acous. Soc. Am. 20, pp. 160-170

Saberi, K. & Perrott, D.R. 1999, *Cognitive Restoration of Reversed Speech,* Nature 398, pp. 760

Whitlock, J.A.T. 2001, An Investigation into the Sensitivity of Intelligibility of "Reversed Segmented Speech" to Acoustical Conditions, Contained in Whitlock, 2003.

Whitlock, J.A.T. 2003, Acoustical Mechanisms Influencing Speech Intelligibility for Primary School Children, Masters' Thesis, University of Auckland

Wilson, O., Dodd, G. et al. 2002, *Classroom Acoustics – A New Zealand Perspective,* The Oticon Foundation, Wellington, New Zealand (ISBN 0-473-08481-3)