IMPROVING THE SPEECH INTELLIGIBILITY IN CLASSROOMS

LAM CHOI LING CORIOLANUS

Ph.D

The Hong Kong Polytechnic University

2010

THE HONG KONG
POLYTECHNIC UNIVERSITY
DEPARTMENT OF
MECHANICAL ENGINEERING

IMPROVING THE SPEECH INTELLIGIBILITY IN CLASSROOMS

LAM CHOI LING CORIOLANUS

A thesis submitted in partial fulfilment of the requirements for
the degree of Doctor of Philosophy

March, 2009

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgement has been made in the text.

____ (Signed)

LAM, Choi Ling Coriolanus (Name of Student)

Abstract

One of the major acoustical concerns in classrooms is the establishment of effective verbal communication between teachers and students. Non-optimal acoustical conditions, resulting in reduced verbal communication, can cause two main problems. First, they can lead to reduce learning efficiency. Second, they can also cause fatigue, stress, vocal strain and health problems, such as headaches and sore throats, among teachers who are forced to compensate for poor acoustical conditions by raising their voices. Besides, inadequate acoustical conditions can induce the usage of public address system. Improper usage of such amplifiers or loudspeakers can lead to impairment of students' hearing systems. The social costs of poor classroom acoustics will be large to impair the learning of children. This invisible problem has far reaching implications for learning, but is easily solved.

Many researches have been carried out that they have accurately and concisely summarized the research findings on classrooms acoustics. Though, there is still a number of challenging questions remaining unanswered. Most objective indices for speech

intelligibility are essentially based on studies of western languages. Even several studies of tonal languages as Mandarin have been conducted, there is much less on Cantonese. In this research, measurements have been done in unoccupied rooms to investigate the acoustical parameters and characteristics of the classrooms. The speech intelligibility tests, which based on English, Mandarin and Cantonese, and the survey were carried out on students aged from 5 years old to 22 years old. It aims to investigate the differences in intelligibility between English, Mandarin and Cantonese of the classrooms in Hong Kong. The significance on speech transmission index (STI) related to Phonetically Balanced (PB) word scores will further be developed. Together with developed empirical relationship between the speech intelligibility in classrooms with the variations of the reverberation time, the indoor ambient noise (or background noise level), the signal-to-noise ratio, and the speech transmission index, it aims to establish a guideline for improving the speech intelligibility in classrooms for any countries and any environmental conditions.

The study showed that the acoustical conditions of most of the measured classrooms in Hong Kong are unsatisfactory. The selection of materials inside a classroom is important for improving speech intelligibility at design stage, especially the acoustics ceiling, to shorten the reverberation time inside the classroom. The signal-to-noise should be higher than 11dB(A) for over 70% of speech perception, either tonal or non-tonal languages, without the usage of address system. The unexpected results bring out a call to revise the standard design and to devise acceptable standards for classrooms in Hong Kong. It is also demonstrated a method for assessment on the classroom in other cities with similar environmental conditions.

Publications arising from the thesis

Conference Paper

1. LAM, CORIOLANUS C. L., LI, K. M. and CHEUNG, STANLEY M. L., “Articulation tests of the first and the second languages in classrooms,” J. Acoust. Soc. Am. **116**, pp. 2611, 2004.

2. LI, K. M. and LAM, CORIOLANUS C. L., “Acoustical conditions of typical classrooms in Hong Kong,” J. Acoust. Soc. Am. **117**, pp. 2405, 2005.

3. LAM, CORIOLANUS C. L. and LEUNG, C. W. (with LI, K. M.), “Comparison of speech communication of renovated classroom by subjective method,” J. Acoust. Soc. Am. **120**, pp. 3199, 2006.

4. LAM, CORIOLANUS C. L. and LEUNG, C. W. (with LI, K. M.), “Case study: Predict acoustical quality of renovated classroom,” J. Acoust. Soc. Am. **120**, pp. 3200, 2006.

Acknowledgements

I am appreciated on the research described in this thesis is financed by The Research Grants Council of The Hong Kong Special Administrative Region. I sincerely, in acknowledgement of the chief supervisor Prof. C. W. Leung and the co-supervisor Prof. K M Li, the academic, technical and administrative support offered by Department of Mechanical Engineering of The Hong Kong Polytechnic University, on the technical and practical parts of the research. I wish to thank Dr S. W. Tang from Department of Chinese and Bilingual Studies of The Hong Kong Polytechnic University, for his unselfish share of the knowledge and information on Cantonese study. Special thanks to Dr W. K. Leung and Dr H. Y. Wong for their knowledgeable inspiration and guidance during my full time study of the research. My dedication of this thesis to my fiancé, my family and my friends, for their encouragement throughout the research.

Table of Contents

Chapter 1
Introduction

1.1 Overview

1.1.1 Background of research

Since last century, the effect of indoor acoustical environments on people has become a controversial topic. As the global population growth and development of countries, the significance of effect of indoor acoustical environments on people is increased. In high rise cities, human voice and speech is often used to provide information. Failure to understand the content of message or low perception of speech led to many social problems. Misunderstanding between people will increase; accidents due to unclear communication at work will happen more easily such as low perception of public announcement in terminal; workforce will become inefficient as well; mental illness will be increase in number due to annoyance; and education of new generation will be retrogressive.

Today, education is the path through which new immigrants and native-born citizens are integrated into the daily life like United States of America, Canada, Australia, and Hong Kong. An educated workforce is a necessary condition for economic growth in today's competitive economy and the current trend towards globalization. Today, in Hong Kong, from data collected by Census Department of the government of Hong Kong, about 22% of Hong Kong's population, a total of about 1.8 million students, is at school or some other form of higher education institution. In the 2003/2004 fiscal year, the Government's total expenditure was HKD 207 billion, and the public spending on education, representing 23.8% of the Government's total recurrent expenditure.

Education is the engine of economic growth; however, the effect of indoor acoustical environments on people is unquestionable. The impact of non-optimal acoustical environment in classrooms is detrimental not only to students but also teachers. The non-optimal acoustical environments in classrooms result in compromised verbal communication and learning. Students will be easily annoyed and frustrated. The achievement and behaviour will be influenced by these factors. Teachers also strain their voice regularly to ensure the students hear it clearly which can often be a cause of claimed ‘industrial injury’ for throats problem. Usually, the inadequate verbal communication leads to the usage of amplifying systems. Those amplifiers or speakers can impair the students’ hearing systems, especially the children. The impact is more significant on young children, hearing impaired and second language students who are integrated into mainstream of classes. It is stringent and crucial that improving the acoustical environment is cost effective.

In fact, one of the main education policies in Hong Kong is to enable students to learn effectively, and to be bi-literate and trilingual. To obtain high speech intelligibility in classrooms for normal-hearing adults working in their mother tongue, researchers have suggested that the reverberation time must not exceed 0.5s to 0.7s [4]; with this reverberation time, the signal-to-noise ratio must exceed 15 dB. One thing that needs to be considered is that adults’ average result on speech intelligibility test is roughly 10% better than young children.

Even more stringent conditions are required in the case of more acoustically challenged persons (young children, the hearing-impaired and second-language speakers). To enable students to learn effectively, it is essential to eliminate the acoustical barriers to learning in the classrooms. Excessive noise and reverberation interfere with speech intelligibility, resulting in reduced understanding and therefore reduced learning. Researchers have also found that excessive background noise and reverberation within a classroom can influence not only speech intelligibility, but also the achievement and behaviour of the students. Insufficient speech levels and long reverberant time may create selective acoustic barriers for hearing impaired children and learning disabled children. As the background noise level increases, teachers have to raise their voice, thus making the students restless and uneasy. As the speech intelligibility within a classroom lowers, both the teacher and the student require greater attention.

Very poor classroom acoustics are obvious to anyone, but the problems caused by marginal classroom acoustics are insidious because they are not universally apparent. Students, teachers and school administrators may not recognize the source of the learning difficulty. Vulnerable students in marginal classrooms may observe that other students understand what teacher was said while they do not. Often, such students are not placed in separate classrooms with enhanced acoustics, but are mainstreamed with other students. These students may develop an unduly low opinion of their cognitive abilities, leading to discouragement, boredom, discipline problems, and possible failure. Improving marginal classroom acoustics will not only help vulnerable students but also benefit teachers and other students by making the classroom a more effective and better learning environment.

1.1.2 Research objectives

Given these considerations, it is clear that a wide range of students benefit from improved classroom acoustics. Better classroom acoustics will help students and teachers, and are a good investment for any country. Thus, this study serves several purposes and objectives.

In this study, Hong Kong serves as an example and a subject, by measuring the acoustics of a range of existing classrooms to compare with the national and international standards for the acoustical design of classrooms for different countries worldwide. With those acoustical parameters measured, establishing an empirical relationship between the reverberation time, background noise level, signal-to-noise ratio, speech transmission index and speech intelligibility. These aim to establish guidelines with practical criteria for assessing the speech intelligibility of classrooms in the design stage, and improving the speech intelligibility. English, Mandarin and Cantonese will be used for the measurement of speech metrics since lack of relevant information in comparison of this area. All these three languages are official languages as media in the classrooms in Hong Kong.

By examine of such classrooms embedded with temporarily acoustic treatment for the enhancement of speech intelligibility in classrooms, acoustics environment will be validated and estimated by means of a commercial software for design stage of the classrooms during urban planning. Data modelling will be developed to predict speech metrics after classroom simulation with varied schemes such as use of closets, shelves/fixtures for breaking up strictly parallel side and front/back walls. With application of engineering know-how on data modelling, speech intelligibility by subjective listening scores with variation on speech metrics that are most influencing to speech transmission will be developed. Field experiments will be conducted to validate the theoretical and numerical relationships.

With proper prediction model and estimation on environment on design stage of classrooms in urban planning, it helps to secure huge investment on the education, both in the schools building and the health of teachers and students. The study illustrates a procedure and setups the guidelines which are applicable to any country and aim for benefit of the society.

1.2 Literature review

1.2.1 Acoustical barriers to learning in classrooms

The adverse effect of noise on classroom activities was recognized as early as the 60s. Since 70s, researches have been carried out to identify acoustical parameters that affect an effective verbal communication in classrooms. The more understanding on relationship of environmental factors to speech transmission, the earlier precaution and measures to plan on design stage of classrooms. Three main environmental factors influence how well students will be able to hear and understand what a teacher says in the classroom: distance from the teacher, reverberation and background noise, include mechanical equipment's noise.

1.2.1.1 Distance from the source of speech

The source of speech, typically teacher's mouth, is considered as a point source. One of the characteristics of sound in point source is decrease of sound energy with distance increase. Crandell and Bess examined the speech recognition of children aged 5 to 7 with normal hearing in a typical classroom environment, with the teacher's speech about 6 dB louder than the background noise [1]. The findings have revealed that the children recognized 89% of the words at 6 feet; 55% at 12 feet; and only 36% at 24 feet. It suggested that children with normal hearing seated in the middle to rear of a typical classroom have more difficulty understanding speech than has traditionally been suspected. Pearsons, Bennett and Fidell measured long-term speech levels at about 2m in front of instructors in high schools [2]. They also measured speech levels at 1m in front of the speakers and various vocal efforts. Most rooms tested by Hodgson in University of British Columbia were "very good" or "excellent" at the front, and "good" or "very good" at the back [3].

1.2.1.2 Reverberation and reverberation time

The second factor is the effect of reverberation in classrooms. Reverberation is the remainder of the sound that exists in a room after the source has stopped. Reverberation time has long been the major objective parameter relating to acoustical conditions in a room. Generally, reverberation time RT60 is the time required for sound energy decayed for 60 dB. It is a useful descriptor of the mean properties of the room and is easily estimated from knowledge of the room materials and geometry. The reverberation in a room generally increases with the room size, decreases with the amount of sound absorption in the room, and is affected by the room shape and contents.

Ideally, reverberation time in classrooms should be in the range of 0.4 – 0.6 seconds. Though, in some cases, the classrooms such as large lecture rooms should have longer reverberation [4]. It is also known that reverberation affects speech intelligibility by affecting the early- and late-arriving energies and, more importantly, their ratio, the early-to-late energy fraction [5]. Children with hearing, processing, language, or attention problems have even more difficulties in reverberant classrooms. Acoustical design criteria for classrooms must include limits on interfering background noise as well as the acoustical properties of the room, which have conventionally been considered in terms of reverberation time [6].

1.2.1.3 Background noise level

The last factor is the background noise levels in classrooms. Background noise masks the spoken word. Masking inhibits our ability to hear one sound in the presence of another sound at the same level, reducing the acoustic cues available to both students and teachers. Students both struggle to hear or become distracted and stop paying attention.

There are many sources of background noise. External noise comes from transportation, construction, school playgrounds, and outer mechanical equipment such as air-conditioners. Noise also comes from spaces surrounding the classroom, such as the cafeteria, gymnasium, corridors and adjacent classrooms. Noise inside the classroom is generated by students activities, chairs scraping, students talking, heating, ventilating and air-conditioning systems and noise generated by the fan of a projector. Besides, masking noise or echo may create by the geometrical configuration inside the classroom.

Today's heating, ventilating and air-conditioning systems use higher, more audible air speeds, and have narrow ceiling ducts that radiate noise. A good guideline is that the noise level in classrooms should not exceed noise criteria NC 25 to 30. Another useful guideline from Seep *et al.* is that the noise level should not exceed 35 dB(A) [4]. Beranek suggested noise criteria NC 30 – 40 for classrooms [7]. Parkin and Humphreys gave an octave noise spectrum that approximates noise criteria NC 25 [8]. Noise criterion NC is widely applicable to access the noise level of mechanical equipments for many years. It is based on empirical corrections to equal loudness contours as standard contours of mechanical noise annoyance to human hearing systems.

Suggestion made on average a critical noise level of 42 dB(A) inside a classroom due to road traffic would affect speech intelligibility [9]. Wang, Shi and Zhong suggested a noise limit of L_{eq} of 50 dB(A) for the background in classrooms [10]. They also found quantitatively that noise not only distracts students but also causes their annoyance and fatigue. It is of interest to point out that they suggested a higher permissible noise criterion for classrooms in China than most other western countries. Hodgson, Rempel and Kennedy reported on the background noise in classrooms that it is comprised of two components: ventilation noise and study activity address [11].

1.2.2 Speech intelligibility in classrooms

The key acoustical characteristic and functional requirement of a classroom is speech intelligibility: students and teacher should be able to communicate effectively without exerted effort. Speech intelligibility means degree or percentage of the speech which the receiver hears it clearly. This is known to be mainly determined by the signal-to-noise level difference – the difference between the speech-signal and background noise levels at a receiver, and the amount of reverberation [6].

Speech intelligibility is directly related to signal-to-noise level difference and is inversely related to the reverberation time. However, in rooms the situation is complicated by the fact that reverberation and steady state levels are interrelated. Excessive reverberation in the classroom can lead to low speech levels, especially at the back of the room [12]. When speech intelligibility is the concern, the amount of reverberation is best quantified by the early-to-late energy ratio [13]. Increased reverberation, while decreasing the ratio of early-to-late energy ratio to the detriment of speech intelligibility has the additional effect of increasing steady state levels by increasing the reverberant sound energy to the benefit of speech intelligibility [14]. Nevertheless, it is more usual to characterize the amount of reverberation in a room by the reverberation time.

Several studies found speech intelligibility scores increased continuously as reverberation time was decreased to zero seconds [15 – 16]. Nábělek again showed that speech intelligibility scores increased continuously as reverberation time was decreased to zero seconds and also that the detrimental effects of reverberation depended on the age of the subject. They also found that students in 10 years old required an extra 5 dB (A) of speech signal level to produce equivalent scores to young adults. Heerwagen and Khiati have reported their assessments of speech intelligibility in classrooms [17]. A survey has highlighted a catalogue of 75 problematic classrooms out of the 400 classrooms on the campus of the University of Washington. Acoustical measurements have also shown high background noise levels and excessive reverberation times for some tested classrooms.

Numerous measurements of the typical classroom were conducted for speech and background noise levels, and for corresponding signal-to-noise ratios. Signal-to-noise is the most influential acoustics parameter to speech intelligibility. It is the indicator to show the sound level difference of signal or speech to background noise at the receiver point. All kind of noise which happened and affected the receiver point are inclusive, such as noise from mechanical equipments. Elliot reported that 7 years old required an extra 5 dB (A) of signal-to-noise ratio to produce equivalent scores to those of young adults [25]. Hougast obtained both speech intelligibility scores and estimates of the speech and background levels in classrooms [26]. He related aggregate speech intelligibility scores to A-weighted signal-to-noise ratios and concluded that a 15 dB (A) signal-to-noise ratio eliminated the detrimental effects of interfering noise.

Researches on speech intelligibility for English and Mandarin have been conducted. It is envisaged taking into account experiences in mainland China and abroad. Those classrooms in primary and secondary schools in Hong Kong would have similar acoustical problems [18 – 20]. It is irrespective of whether they are fitted with air-conditioners or not. In addition, speech intelligibility tests were carried out mainly for English although there have been some recent studies that investigated the effect of intruding noise in classrooms where Mandarin was the principal language of communication. Nevertheless, in Mandarin there have been investigations on speech intelligibility [21 – 22]. Experimental results have indicated that Mandarin is comprised of monosyllabic characters which is able to tolerate a longer reverberation time than English for a given speech intelligibility.

If speech intelligibility in an enclosure is satisfactory for English, it is not necessarily satisfactory for Chinese, or vice versa. The results suggest that in terms of speech intelligibility, Mandarin is slightly better than English under reverberant conditions, and English is considerably better than Mandarin under noisy conditions. In both

English and Mandarin there have been considerable investigations on speech intelligibility [21, 23 – 24]. However, it appears to be inappropriate to directly compare these relationships between the two languages because they were established under different conditions. Moreover, these relationships were mainly for diffuse field.

Bradley conducted speech intelligibility tests and acoustical measurements for background noise, early decay times, reverberation times and speech transmission index in classrooms for secondary school students in Canada [6]. School children with no reported impaired hearing were used as subjects for the tests to obtain the speech intelligibility scores. It was demonstrated that Speech Transmission Index (STI) was one of the most accurate predictors of speech intelligibility scores.

1.2.3 Physical measures for the quality of speech transmission

The common identifiers of the efficiency of speech transmission, or speech transmission quality have been studied. In the past decades, there have been considerable investigations on speech intelligibility.

Speech intelligibility is the measure of the effectiveness of speech. The measurement is expressed as a percentage of a message that is understood correctly. However, speech intelligibility does not implied speech quality. A synthesized voice message may be completely understood by the listener, but judged to be harsh, unnatural, and of low quality. A message that lacks quality still be intelligible. Many objective indices, predictors and methods have been developed and established. Different types of acoustical measures were compared as predictors of speech intelligibility of varied sizes and acoustical conditions in the past decades. These included signal-to-noise ratio (SNR), reverberation time (RT), early-to-late sound ratios (C) and useful-to-detrimental Ratios (U), articulation index (AI), and speech transmission index (STI).

1.2.3.1 Signal-to-Noise Ratio (SNR)

Signal-to-noise ratio (SNR), or speech-to-noise signal in this study, is the ratio of energy received from signal to energy received from background noise. The background noise creates a masking or disturbance to interfere the signal at the receiver. SNR is in a dynamic range which express in logarithmic decibels.

$$SNR = 10\log\left(\frac{P_S}{P_N}\right)$$

$$= 10\log P_S - 10\log P_S$$

where P_S is sound pressure of signal or speech and P_N is sound pressure of background noise,

$$SNR = S - N$$

where S is signal noise level or speech level and N is background noise level.

Signal-to-noise ratio SNR is the most significant factor to speech intelligibility. Definitions of SNR have characteristics to include all masking or intruding noise into calculation as shown in the above equations. It is the sound level difference between signal and background noise. The background noise level is summation of energies of all kind of noise such as transportation, playground activities, mechanical equipment and echo noise induced by reverberation. Since consideration of signal energy and background noise energy in sound pressure, it is further evaluated and related to speech intelligibility metrics. Lochner and Burger found SNR is parameter which clearly reflects the effect of reverberant and so forth to speech intelligibility [37, 40]. They conclude that the early arrived energy of reverberant noise does not take into account to the signal noise energy at the beginning but increases the measured speech transmission index (which will be discussed later in this chapter). As the late arrived energy of reverberant noise takes effect to arrive to the receiver, it masks the signal and the measured speech transmission index decreases. In consideration of those effects, SNR is more reliable and most considerable factor to speech intelligibility.

1.2.3.2 Reverberation time (RT)

Reverberation time (RT) is the time required for sound energy of the signal in a room to decrease by certain decibels after the signal stops. Typically RT is in RT 60 or the shorter form RT 30 or RT 10. RT 60 is the time required for sound energy decayed by 60 dB and the same to RT 30 and RT 10. Though, RT 60 is widely used for measurement and RT 10 is measured but need to extrapolate for longer period. The larger RT value indicates the most reverberant of the field.

For a point source, like speech in a reverberant field which reverberation takes effect inside such as a room, the energy of the sound have relationship with reverberation time:

$$L_W = L_P + 10\log\frac{V}{RT} - 14dB$$

where L_W is sound power level, L_P is sound pressure level, V is volume of the room and RT is reverberation time of the room.

Many researches have been conducted to propose different formulas for predicting reverberation times [34 – 36]. Most of the proposed formulas are based on sound diffuse field assumption in calculation. Reverberation time RT can be presented into a semi-empirical equation that relate the reverberation time of a live room to the room volume and the properties of the room boundaries

$$RT = \frac{24\ln 10}{Ac}$$

where A is the room absorption and c is the speed of sound in the room.

In addition to the ray-based prediction programs like RAYNOISE, there are two simple formulas commonly used to predict reverberation time in a fully enclosed space: Sabine equation and Eyring reverberation formula. These two formulas are slightly different because they are derived from somewhat different considerations [32 – 33].

Sabine equation:

$$A = S\overline{\alpha}$$

$$S\overline{\alpha} = \sum \alpha_i S_i$$

where S is the total surface area of the room boundaries and $\overline{\alpha}$ is the average value of the statistical absorption coefficient.

Eyring equation:

$$A = S \ln\left(\frac{1}{1-\overline{\alpha}}\right)$$

where S is the total surface area of the room boundaries and $\overline{\alpha}$ is the average value of the statistical absorption coefficient.

However, in classrooms, the sound pressure at distance r away from a point source is more accurately described by

$$P_e^2 = \rho_o c_o \overline{W}\left(\frac{Q}{4\pi r^2} + \frac{4}{R}\right)$$

where

$$R = \frac{S\overline{\alpha}}{1-\overline{\alpha}}$$

where $\overline{W}$ is the sound power of the source, Q is the directivity factor due to boundary effect or source directivity, and $\overline{\alpha}$ is the sound absorption coefficient.

As shown in the equations above, RT is an acoustics parameter highly dependent on the room size and absorption inside the room. A larger room have longer reverberation time, more reverberant of a room and poorer the speech intelligibility inside a room. Longer reverberation time determines the longer time for sound energy to decrease. However, sound propagation in air is not changed which means that sound energy is reflecting from walls or obstacles inside a room without loss or absorb. Besides, if the sound power energy or sound pressure energy are well known with the volume of the room, RT of the room can be estimated. It is also known that reverberation is about one of the energies´ characteristics inside a room. The early- and late-arriving energies implicate how much energy is absorbed inside a room after reverberation, in terms of early-to-late energy ratio. Change of early-to-late energy ratio leads change in reverberation time and in terms affect the speech intelligibility [5, 37].

1.2.3.3 Early-to-late sound ratio (C) and useful-to-detrimental ratio (U)

Early-to-late ratio (C_n) or Klarheitsmass is the ratio of the early arriving sound energy in the first n second after the direct sound to the late sound energy arriving after n second [35]. Alim researched and modified based on Thiele findings on definition which larger value of definition should have higher speech intelligibility [39]. He defined and concluded that clarity is more applicable than definition for music [41]. For n in 50ms, many researches shown at about 50ms, early-arriving energy enriches the speech intelligibility but interfere the speech after 50ms. It has been shown to relate to the speech intelligibility scores [37]

$$C_n = 10\log\left(\frac{\int_0^n p^2(t)dt}{\int_n^\infty p^2(t)dt}\right)$$

where the time for early-arriving energy is from 0 second to n seconds and the time for late-arriving energy is after n seconds.

Bradley have been studied that it is safer and more generally reliable to use the n second useful-to-detrimental ratio (U_n) as the preferred predictor of speech intelligibility [5]. This has the added advantages that C_n values from which U_n values are derived are useful in rooms where not only to perform music. Without consideration of background noise energy, he found U_n can simplifies into C_n when n equals to 50ms. The use of U_n values based on early-to-late ratios is also desirable as these values are easily related to the fundamental physical quantities involved

$$U_n = \frac{C_n}{1 + (C_n + 1)\dfrac{E_{BNL}}{E_{SL}}}$$

where E_{BNL} is related background noise energies and E_{SL} is the related speech energies.

It is also known that reverberation affects speech intelligibility by affecting the early- and late-arriving energies and their ratio of the early-to-late energy fraction, which is more important [5, 37].

1.2.3.4 Articulation index (AI)

Articulation index (AI) is firstly introduced by French and Steinberg, later modified by Kryter, to express the quality of speech transmission [42 – 43]. It is a common and the simplest method for determining the quality of speech transmission channels by performing intelligibility tests with talkers and listeners using sentences, rhythm words, or other test material [27 – 28]. It was developed mainly to account for distortions in the frequency domain (noise, band pass limiting). The measure is based on steady state signal-to-noise concepts. The AI method is particularly appropriate for channels with distortion in the frequency domain such as interfering noise and band pass limiting.

The calculation of the articulation index basically consists of three steps [29]:

1. The calculation of the effective signal-to-noise ratio within a number of frequency bands. For each band the signal level and the noise level are considered, together with auditory masking introduced by the hearing organ.
2. A linear transformation of the effective signal-to-noise ratio to an octave band specific contribution to the AI, ranging from one to zero. This transformation is such that, for each band, the maximum contribution of one is reached at a signal-to-noise ratio of +18 dB and the minimum contribution of zero is reached at a signal-to-noise ratio of −12 dB.
3. The calculation of the weighted mean of the contributions of all relevant octave bands constitutes the AI.

1.2.3.5 Speech transmission index (STI)

Speech transmission index (STI) was firstly derived from modulation transfer functions in the 70s by Houtgast and Steeneken [29 – 30] and has recently been demonstrated to relate to speech intelligibility test scores in rooms [5, 29 – 30]. It has been proposed based on modulation transfer functions and including the effects of interfering background noise. It indicates the speech transmission from modulation transfer functions of sound path from the signal to the receiver. Modulation transfer functions (MTF) is change of modulation frequencies on averaged frequency band of the signal at the receiver. With conversion of MTF into a functional index, it is speech transmission index STI.

Speech transmission in a room is in impulsive response. With characteristics of impulse response, there have a modulation index for every modulation frequency under the modulation transfer function of the system

$$m(F) = \frac{\left| \int_0^\infty e^{-j2\pi Ft} p^2(t)dt \right|}{\int_0^\infty p^2(t)dt}$$

where p(t) is function of impulse response of the signal.

When considering the presence of background noise, the modulation becomes

$$m(F) = \frac{\left| \int_0^\infty e^{-j2\pi Ft} p^2(t)dt \right|}{\int_0^\infty p^2(t)dt} (1 + 10^{(-\frac{SNR}{10})})^{-1}$$

where SNR is signal-to-noise ratio at the receiver.

Typically reverberation is taken into account on the situation, the modulation then becomes

$$m(F) = \left[1 + (\frac{2\pi FT}{13.8})^2 \right]^{-\frac{1}{2}} (1 + 10^{(-\frac{SNR}{10})})^{-1}$$

where F is modulation frequency of the signal and T is reverberation time of the room.

When the reverberation time becomes increasingly large in relation to the period of the modulation, the depth of the modulation is increasingly reduced by the acoustical properties of the room. Since the modulation is decreased with existence of both reverberation time and background noise, SNR is used to simplify the effect of these two parameters that SNR is dependent on variation of reverberation time and background noise. Therefore, the function becomes

$$SNR_{app} = 10 \log \frac{m(F)}{1 - m(F)}$$

where SNR_{app} is apparent signal-to-noise ratio.

In the frequency band 125 Hz to 8k Hz, 7 modulations acquires with analysis from its one-third octave frequency response. With calculation on matrix of those modulations, an apparent SNR is acquired. Apparent SNR is limited between +15 dB as SNRapp = +15 dB for SNRapp > 15 dB or < 15 dB since SNR is comparatively high which negligible in the modulation transfer function. As the apparent SNR is originally computed from different frequency band, the equation then converted into a single parameter for averaging as

$$\overline{SNR}_{app} = \sum_{k=1}^{7} w_k (SNR_{app})_k$$

where w_k is frequency weighting constant in 125 Hz to 8k Hz (0.129, 0.143, 0.114, 0.114, 0.186, 0.171 and 0.143 respectively).

With this single parameter, the speech transmission index from the modulation is then concluded into an index as

$$STI = \frac{\overline{SNR}_{app} + 15}{30}$$

The most important of STI is representation into SNR related parameter, SNRapp. It is clearly stated and emphasizes the importance of SNR to assess and improve the speech intelligibility. With dominant speech modulation frequencies in the certain range of frequencies, there would be some reduction in the modulation depth in almost all real rooms. The reductions of the amplitude modulation of test signals at various acoustical frequencies as a function of speech modulation rate are combined to produce this one index STI [31].

1.3 Research methodology

The aim of this study is to improve speech intelligibility within existing classrooms. Previous researches conducted abroad have shown the acoustics of many classrooms in general to be inadequate, with poor speech intelligibility due to the reverberation time being too long, coupled with excessive background noise. Classroom acoustics have been investigated in a number of overseas countries but a few, if any, investigations have been undertaken in schools in Asia.

This study firstly gathers and reviews information and literatures on classrooms acoustics. To carry out the subjective intelligibility tests, we should have our database of languages in Cantonese which is one of the media of the speech since Hong Kong is officially in tri-lingual as education media in the schools. A procedure on development of database of phonetically balanced word lists is setup as reference for other educators in different languages. Phonetically balanced word lists in Cantonese are constructed for subjective intelligibility tests.

After the preparation of testing materials, the national and international standards and guidelines in classrooms are studied for comparison on acoustical measurements and surveys on 20 selected classrooms with similar geometrical characteristics but different acoustics features in Hong Kong are carried out. Acoustical parameters such as indoor ambient noise, outdoor noise, noise criteria from mechanical ventilation systems, reverberation time, speech transmission index, and rapid speech transmission index are studied and compared with those standards for more understanding on the situation in Hong Kong.

One of the objectives on this study is to predict the acoustical performance inside a classroom for preliminary design stage of classroom during urban planning. By using the computer simulation and modelling software, RAYNOISE to adjust and apply different types of acoustical measures and treatment inside a classroom, the empirical consideration as guidelines for classroom is defined.

Though the prediction by ray-based computer program is useful as reference, the main purpose on this study is to improve the speech intelligibility. Speech intelligibility is subjective feeling to human hearing. It also depends on their knowledge and experiences growing with their ages. The subjective intelligibility tests are carried out in 5 selected classrooms with students to be the subject who’s aged from 5 years old to 22 years old, by use of three different languages as testing media: English, Mandarin, and Cantonese. After the listening tests, word scores or PB scores is calculated and empirical relationships are found with acoustical parameters. Although many acoustical parameters have an influential effect on speech intelligibility, this study is focused on the effect of signal-to-noise ratio, which has the strongest impact on speech intelligibility; reverberation time, which can be defined at the preliminary design stage; speech transmission index, which is widely use as objective indicator to speech transmission; to PB scores, which reflects exactly how much of the content of the speech is received. Besides, factors such as age and languages´ characteristics are compared for understanding of various educational needs.

A data modelling by neural network is created and used to predict the effect of PB scores in variation of objective acoustical parameters: signal-to-noise ratio, reverberation time, and speech transmission index. With the help of this engineering know-how as tool for prediction and estimation, program or web-based program can be easily created for every educators and designers to plan for their schools and classrooms design.

Guidelines and empirical relationships are expected to establish and conclude for improving the speech intelligibility in classrooms. It also forms a basis for the development of a design guideline suggesting methods of improving the acoustics within existing classrooms in Hong Kong. Certain ranges in those acoustical parameters should be maintained for newly designed classrooms, in compromise of huge loss to investment on healthy and knowledge standards of citizens. Stringent criteria should be concerned for children, hearing impaired students or learning difficulties students.

References

1. CRADELL, C. and BESS, F., “Speech recognition of children in a typical classroom setting”, ASHA **29**, pp. 87, 1986.

2. PEARSONS, K. S., BENNETT, R. L. and FIDELL, S., “Speech levels in various noise environment”, Bolt Beranek and Newman Inc., Report to the USEPA, Canoga Park, CA, May 1977.

3. HODGSON, M., “Rating, ranking and understanding acoustical quality in university classrooms”, J. Acoust. Soc. Am. **112**, pp. 568-575, 2002.

4. SEEP, B., GLOSEMEYER, R., HULCE, E., LINN, M. and AYTAR, P., “Classroom Acoustics Booklet”, Published by the Acoustical Society of America, August 2000, http://asa.aip.org/classroom/booklet.html.

5. BRADLEY, J. S., “Predictors of speech intelligibility in rooms”, J. Acoust. Soc. Am. **80**, pp. 837-845, 1986.

6. BRADLEY, J. S., “Speech intelligibility studies in classroom”, J. Acoust. Soc. Am. **80**, pp. 846-854, 1986.

7. BERANEK, L. L., “Noise and Vibration Control”, McGraw Hill, New York, 1971.

8. PARKIN, P. H. and HUMPHREYS, H. R., "Acoustics, Noise and Buildings", Faber, London.

9. HOUGAST, T., "The effect of ambient noise on speech intelligibility in classrooms", Applied Acoustics **14**, pp. 15-25, 1981.

10. WANG, J. Q., SHI, J. H. and ZHONG, X. Z., "Effects of traffic noise on pupil behaviour in classroom", Acta Acoustics **17**, pp. 248-255, 1992.

11. HODGSON, M., REMPEL, R. and KENNEDY, S., "Measurement and prediction of typical speech and background noise levels in university classrooms during lectures", J. Acoust. Soc. Am. **105**, pp. 226-233, 1999.

12. HODGSON, M., "Experimental investigation of the acoustical characteristics of university classrooms", J. Acoust. Soc. Am. **106**, pp. 1810-1819, 1999.

13. BISTAFA, S. R. and BRADLEY, J. S., "Reverberation time and maximum background noise level for classrooms from a comparative study of speech intelligibility metrics", J. Acoust. Soc. Am. **107**, pp. 861-875, 2000.

14. HODGSON, M. R. and NOSAL, E. M., “Effect of noise and occupancy on optimum reverberation times for speech intelligibility in classrooms”, J. Acoust. Soc. Am. **111**, pp. 931-939, 2002.

15. NÁBĚLEK, A. K. and PICKETT, J. M., “Reception of consonants in a classroom as affected by monaural and binaural listening, noise, reverberation, and hearing aids”, J. Acoust. Soc. Am. **56**, pp. 628-639, 1974.

16. NÁBĚLEK, A. K. and ROBINSON, P. K., “Monaural and binaural speech perception in reverberation for listeners of various ages”, J. Acoust. Soc. Am. **71**, pp. 1242-1248, 1982.

17. HEERWAGEN, D. and KHIATI, T., “Assessing speech intelligibility in classrooms at the University of Washington”, Proceeding of ICA **1998**, pp. 2729-2730, 1998.

18. CHE, S. K. and KANG, J., “Criteria and permissible limit for noise in classrooms of primary and middle school”, Proceeding of WestPrac **II**, Hong Kong, pp. 24-29, 1985.

19. WANG, J. Q. and GU, Q. G., “Classroom acoustic criteria in school building regulations”, Proceeding of Internoise **87**, pp.

1141-1144, 1987.

20. GOLD, M. A. et al, “Classroom acoustics II: Acoustical conditions in elementary school classrooms”, Proceeding of ICA **1998**, pp. 2723-2724, 1998.

21. ZHANG, J. L., “Speech” in “Handbook of Acoustics”, edited by MAA, D.Y. and SHEN, H., Science Press, Beijing, Ch. 19, pp. 404-435, 1987.

22. KANG, J., “Comparison of speech intelligibility between English and Chinese”, J. Acoust. Soc. Am **103**, pp. 1213-1216, 1998.

23. BERANEK, L. L., “Acoustic Measurements”, Wiley, New York, Ch. 13, pp. 625-634, Ch. 17, pp. 761-792, 1949.

24. HOUTGAST, T. and STEENEKEN, H. J. M., “A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria”, J. Acoust. Soc. Am **77**, pp. 1069-1077, 1985.

25. ELLIOT, L. L., “Effects of noise on perception of speech by children and certain handicapped individuals”, J. Sound Vib. **16**, pp. 10-14, 1982.

26. HOUGAST, T., “The effect of ambient noise on speech

intelligibility in classrooms", Applied Acoustics **14**, pp. 15-25, 1981.

27. ANSI S3.2-1989 (R1999), "Method for measuring the intelligibility of speech over communication systems", American National Standards Institute, New York.

28. ANSI S3.5-1997 (R2002), "Methods for calculation of the speech intelligibility index", American National Standards Institute, New York.

29. STEENEKEN, H. J. M. and HOUTGAST, T., "A physical method for measuring speed-transmission quality", J. Acoust. Soc. Am. **67**, pp. 318-326, 1980.

30. HOUTGAST, T., STEENEKEN, H. J. M. and PLOMP, R., "Predicting speech intelligibility in rooms from the modulation transfer function. I. General room acustics", Acustica **46**, pp. 60-72, 1980.

31. BRADLEY, J. S., "Auditorium acoustics measures from pistol shots", J Acoust. Soc. Am. **80**, pp. 199-205, 1986.

32. SABINE, W. C., "Collected Papers on Acoustics", Peninsula, Los Altos, CA, pp. 43-52, 1922.

33. EYRING, C. F., “Methods of calculating the average coefficient of sound absorption”, J. Acoust. Soc. Am. **4**, pp. 178-192, 1933.

34. MILLINGTON, G., “A modified formula for reverberation”, J. Acoust. Soc. Am. **4**, pp. 69-72, 1932.

35. FITZROY, D., “Reverberation formula which seems to be more accurate with nonuniform distribution of absorption”, J. Acoust. Soc. Am. **31**, pp. 893-997, 1959.

36. ARAU-PUCHADES, H., “”An improved reverberation formula”, Acustica **65**, pp. 163-180, 1988.

37. LOCHNER, J. P. A. and BURGER, J. F., “The influence of reflections on auditorium acoustics”, J. Sound Vib. **1**, pp. 426-454, 1964.

38. STRØM, S., “Concert Hall Acoustics”, University of Trondheim, Norway, ELAB Report STF44, A82006, 1982.

39. THIELE, R., “Richtungsverteilung und Zeitfolge der Schallruckwurfe in Raumen (Directional distribution and time series of acoustic noise reflections in rooms)”, Acustica **3**, pp. 291, 1953.

40. LOCHNER, J. P. A. and BURGER, J. F., “The intelligibility of

speech under reverberant conditions”, Acustica **11**, pp. 195-200, 1961.

41. ALIM, A. *et al.*, “Dependence of time and register definition of room acoustical parameters with music performances”, Dissertation, TU Dresden, 1973.

42. FRENCH, N. R. and STEINBERG, J. C., "Factors governing the intelligibility of speech sounds", J. Acoust. Soc. Am. **19**, pp. 90-119, 1947.

43. KRYTER, K. D., “Methods for the Calculation and Use of the Articulation Index”, J. Acoust. Soc. Am. **34**, pp. 1689-1697, 1962.

44. DANCE, S. M. and SHIELD, B. M., “Modeling of sound fields in enclosed spaces with absorbent room surfaces. Part I: Performance spaces”, Applied Acoustics **58**, pp. 1-18, 1999.

45. DANCE, S. M. and SHIELD, B. M., “Modeling of sound fields in enclosed spaces with absorbent room surfaces. Part II: absorptive panels”, Applied Acoustics **58**, pp. 373-384, 2000.

46. DANCE, S. M. and SHIELD, B. M., “Modeling of sound fields in enclosed spaces with absorbent room surfaces. Part III:

Barriers", Applied Acoustics **58**, pp. 385-397, 2000.

47. BISTAFA, S. R. and BRADLEY, J. S., "Predicting reverberation times in a simulated classroom", J. Acoust. Soc. Am. **108**, pp. 1721-1731, 2000.

48. BISTAFA, S. R. and BRADLEY, J. S., "Predicting speech metrics in a simulated classroom with varied sound absorption", J. Acoust. Soc. Am. **109**, pp. 1474-1482, 2001

Chapter 2
Development of a database for evaluating the speech intelligibility of Cantonese

2.1 Introduction

Speech intelligibility is an acoustic and perceptual characteristic and a functional requirement for assessing the effectiveness of a communications channel. Good speech clarity is required in such places as bus and railway stations, airport terminals, and classrooms. An informative announcement delivered effectively through public address systems in stations and terminals is very important from the aspect of safety. A higher quality of speech intelligibility is needed in classrooms to facilitate the learning process for students. Furthermore, more stringent acoustical conditions are required for more acoustically and linguistically challenged persons such as young children, the hearing impaired, and those listening to a second language. By the need to improve the acoustical qualities of classrooms for primary and secondary schools in Hong Kong, a subjective evaluation is required to observe the effects on communication in classrooms due to background noise, reverberations, and other factors. A subjective intelligibility test based on the students' mother tongue, Cantonese in this case, is needed. To conduct speech intelligibility tests, a list of phonetically balanced words in the language is normally required.

We describe the background and justify the need for the current study, which provides detailed procedures for the development of a database that will be used in speech intelligibility tests. Cantonese is used as an example to lay down the general methodology for such a development.

2.2 Background

2.2.1 Languages spoken in the world

How many languages are spoken by the peoples in the world? According to *Ethnologue* [1], there is no exact figure but carefully-compiled records have suggested that there are more than 6,800 different languages, of which over 400 are classified as nearly extinct because of the steady decline in the number of speakers. It is remarkable that half of today's languages have fewer than 10,000 speakers and a quarter have fewer than 1,000. Fifty of these 6,800 languages are thriving. Each of these 50 languages has no fewer than 20 million speakers. Chinese (Mandarin) tops the list with 885 million first language speakers. Spanish and English rank second and third, having 332 and 322 million first language speakers, respectively. Counting second language speakers, English is the most popular language in the world. Language can undeniably be a means of effective communication among peoples. The lack of it may hinder meaningful conversation and cause unnecessary conflicts between groups of people. It is abundantly clear that communication in the first languages or "mother tongues" of the peoples of the world is the most effective method of communicating [2].

2.2.2 Languages in the role to evaluate speech intelligibility

Given the importance of communication, there were many early studies on the development of subjective and objective measures of the quality of speech communication channels. A human perception or articulation test is the most common way to evaluate the intelligibility of speech. Such subjective tests as the Harvard Intelligibility Test of Phonetically Balanced (PB) Words [3], the Fairbanks Rhyme Test [4], the Modified Rhyme Test [5], and the Harvard Test Sentences [3] were developed. There were other studies [6, 7] that investigated the relation between the scores obtained from different intelligibility tests. Generally speaking, there is an advantage of using monosyllabic PB words for listening tests because the result indicates more accurately the phonemes that the subject actually heard [3]. In addition, the list of the PB words has a notable characteristic: the choice of words reflects the relative frequencies of occurrence of phonemes and their distribution in normal speech used daily. English is the main language used in these earlier studies to develop the tests of subjective intelligibility, although Houtgast and Steeneken [8] have developed a word list comprised of phonetically balanced consonant-vowel-consonant words that are predominantly nonsense syllables, which are meaningless on their standalone usage, in the Dutch language.

2.2.3 Physical measures and subjective measures for speech intelligibility

Taking a different approach, French and Steinberg [9] used an articulation index (AI) to evaluate speech intelligibility by using physical measures. Steeneken and Houtgast [10] extended this concept by measuring the effective signal-to-noise ratio in a number of octave bands leading to a physical parameter known as the speech transmission index (STI). In their study, the Dutch language was used for a variety of communication distortions. The speech transmission index was also verified to be an effective physical measure for estimating the speech intelligibility of a communications channel for other languages [11, 12]. A technical device was developed in the 1980s to measure the speech transmission index [13]. A rather similar physical parameter known as Rapid Speech Transmission Index (RASTI) was also developed, in which the required octave bands were reduced to two octave bands at 500 and 2000 Hz. Steeneken and Houtgast [14] used RASTI to evaluate speech intelligibility in a number of different languages as well.

There is an apparent advantage of using a physical parameter, such as the STI, to evaluate the quality of speech communications channels. Its introduction can possibly eliminate the need to conduct carefully controlled listening tests that are both expensive and time-consuming. However, there are major problems in employing the STI and other physical parameters to substitute for standard speech intelligibility tests using listeners. Schmidt-Nielsen [15] pointed out that the standard error of estimate for the STI is too large. More importantly, there is no evidence to support the view that the improvement of STI scores necessarily leads to an enhancement of speech intelligibility or vice versa. In fact, the STI does not reflect the difference in intelligibility scores for different speakers under the same acoustical conditions. Despite these apparent shortcomings in the use of objective tests, the STI has been widely accepted as a method to assess speech intelligibility [16, 17]. Some current developments may lead to improvements in the reliability of the STI in evaluating speech intelligibility in a communications channel, see references 18 and 19.

In these days of ethnic pride and growing appreciation of minority groups, educators, government officials, and missionaries are realizing more than ever the importance of using first languages, as this has been proven to be the most effective means of communicating. It is also considered in deciding on which languages to use for educational and literary purposes, how well the vernacular speakers understand their national language, trade language, or other second language, their ability to use different languages in different social domains, differing levels of bilingual proficiency across a language population, and whether or not they will accept educational materials and other literature in those languages.

In recent years, there has been a resurgence of interest in using various subjective tests to explore speech intelligibility under various acoustical conditions using different languages involving both native and non-native speakers and listeners [12, 20 – 27]. Despite these and earlier efforts, many languages still lack suitable test materials to subjectively evaluate the quality of speech communications. For instance, to the best of our knowledge, there are no appropriate PB word lists for Spanish, Portuguese, Russian, French, Hindi, Bengali, Cantonese (Chinese), and many languages that are ranked among the top 50 spoken in the world. Although one can use the results obtained from English or Dutch, for example, for extrapolation to other languages, doubts have been cast on the general applicability of such an approach because of the known language-specific effects [14]. It is thus desirable to develop different word lists to use in subjective tests of the speech intelligibility of different languages.

2.3 Methodology on developing test materials

2.3.1 Phonetically balanced monosyllabic words

The PB monosyllabic words are by far the most popular test materials for subjectively assessing speech intelligibility because the articulation scores indicate the number of words actually heard by the subject. In fact, the articulation scores for short words or sentences are highly affected by psychological factors, and by knowledge and experiences of the subject. The test results for PB monosyllabic words are comparatively more reliable for speech intelligibility assessment than tests using short words or sentences as they are very sensitive to signal-to-noise ratio [29]. As a result, a more accurate and precise interpretation of the speech intelligibility in a room could be better assessed by articulation scores based on the former material. Hence, we focus on developing PB word lists for conducting speech intelligibility tests.

Indeed, the importance of developing a database of word lists for different languages cannot be overstated. Consider the following example: Australia, where English is the first language, has developed a separate list of Australian English PB words because the Australian accent differs from that of American English [28]. From this, we see that there is a need to develop a database of PB word lists for other languages.

Over the past decades, paradigms based on PB monosyllabic words for English, Mandarin, and Australian English have been developed and published as national standards [29 – 31] for the United States, the People's Republic of China, and Australia, respectively. Unfortunately, information or guidelines for the development of such paradigms for other languages are not easily obtainable as most of these publications only provide a summary of the respective word lists. Without a sufficient knowledge of linguistics, it would be difficult and time consuming for acousticians or other educators to develop such a database for a given language.

The principal objective of this chapter is to provide, for future reference, guidelines for developing a PB word list of a language. To illustrate the methodology, Cantonese is chosen as an example to demonstrate the procedures used to establish a set of PB monosyllabic word lists for the speech intelligibility test. In fact, the development of PB monosyllabic word lists for Cantonese has an interest of its own because about 66 million people [1] use Cantonese as their first language and many others use it as their second language.

2.3.2 Characteristics of Cantonese

Cantonese is one of the major dialects [32] of the Chinese language. Indeed, Cantonese has a history of about 2,000 years, while Mandarin has only been around for about 700 or 800 years. Cantonese is a language spoken in the south eastern region of mainland China, in Hong Kong, and Macau, and by some Chinese minorities throughout Southeast Asia in such places as Singapore, Malaysia, Thailand, Indonesia, and Vietnam. It is not only the mother tongue of most local residents in Hong Kong, but is also a medium of instruction in the local education system. In fact, Cantonese, not

Mandarin, is often the dominant language in overseas Chinese communities. This is largely due to the fact that, in the mid to late 20^{th} century the largest flow of Chinese immigrants to many places of the world was from Hong Kong. There have been several investigations using Cantonese for speech processing and recognition systems [33 – 36], but there has been no systematic study to compile a list of PB Cantonese words to measure the speech intelligibility in a communications channel.

To establish suitable test materials, it is important to select PB words from all monosyllabic words. These chosen words will be grouped into balanced lists, with lists of equal difficulty in each list [3]. As with most other Chinese dialects including Mandarin, Cantonese uses the same set of Chinese characters in writing, although different sets of vocabulary have been developed throughout the historical development of different dialects. Chinese characters are essentially based on logograms, where each symbol represents a morpheme – a meaningful unit of language. Each Chinese character is monosyllabic, and an equivalent English word may be composed of one or more Chinese characters. For example, one Chinese character is needed to translate the English word “horse” and two characters are needed for “apple”. Few words, except those translated from foreign languages, require more than four Chinese characters (or syllables) to represent. According to a survey based on 710,000 words extracted from many different articles in literary, scientific, and political publications [37], Zhang discovered that the percentages of Chinese words with one, two, three, and four syllables were 6.01%, 74.14%, 11.99%, and 7.18%, respectively. Thus, it is perhaps not surprising to find that most Chinese dialects are essentially polysyllabic languages [38]. Further, it is more significant to use ‘syllable’ than ‘word’ to describe for Chinese characters. As the nature of Chinese characters (or syllables) mentioned and Cantonese is also polysyllabic languages, PB monosyllabic word lists is our destination rather than PB word lists. In Cantonese or Mandarin or even most Chinese dialects, word lists can be in one, two, three and four syllables [37]. For PB monosyllabic word lists, it only limited to one syllable word. The more syllables used as test material for speech intelligibility assessment, the lower the sensitivity and reliability of the results, the same applies to the sentences used for the test material.

2.3.3 Corpus of standardized Cantonese

Due to the characteristics of Cantonese mentioned above, our first step was to select sets of PB characters. For other languages, a list of monosyllabic words should first be selected from “standard” literature. For instance, a comprehensive dictionary of the language would be a good choice because it should include most of the vocabularies or single words of that particular language. There is a slight complication in this case because Cantonese uses the same system of characters as Mandarin, but often uses different words that have to be written in different characters. More importantly, Cantonese sounds quite different from Mandarin, mainly because it uses a different set of syllables. It would be rather difficult to find a suitable Chinese dictionary that only covers words spoken in Cantonese because most of these dictionaries are based on Mandarin.

Unlike standard Mandarin, there is no official agency to regulate standard Cantonese. However, it is noteworthy that The Linguistic Society of Hong Kong (LSHK) has compiled a comprehensive table in an attempt to standardize the most frequently used characters that are based on the Cantonese Romanization of Chinese characters [39, 40]. A corpus of 13,053 Chinese characters has been collected in the guide. In the corpus there are 10,676 characters that have a Cantonese pronunciation. As a result, it is reasonable to take this table as a good starting point in constructing lists of PB characters.

Generally speaking, two phonological units, initials and finals, can be classified in standard Cantonese. The initials, sometimes known as the onsets, are initial consonants of all possible syllables. The finals, sometimes known as rhymes, are the part of the syllable that remains after the initials are taken off. There are 19 initials and 61 finals in Cantonese where the null initials are not represented. According to LSHK, both the initials and finals are represented by Romanized alphabets.

The pronunciation of a Cantonese syllable can be constructed by combining different initials and finals. Based on 20 initials and 63 finals (which included null initials and null finals), there are only 667 basic syllables because not all combinations of initials and finals form a valid syllable. Almost all Chinese dialects including Cantonese use tones. A character is distinguished not only by its combination of initials and finals but also by the tone in which it is pronounced. The standard Cantonese has nine tones in six distinct tone contours. To simplify the system, LSHK has proposed to use just six numeric markers to differentiate these distinct tone contours only. It is noteworthy that tone is an important element in Chinese languages [37, 41]. Each tone mark is placed at the end of the syllable, for example, hoeng1 gong2 (Hong Kong). If the tone is taken into account, there are about 1,700 valid tonal syllables. Given that the total number of Chinese characters is well over 12,000, it is not uncommon to find some Chinese characters that are pronounced with exactly the same syllables and tones.

There is also another interesting feature about Chinese characters that is worth mentioning. Some characters have two or more pronunciations; i.e., the same character may be uttered with different syllables and tones. Some of these characters may have different meanings and sound different. Others may have the same meaning but sound different. These groups of characters are known as ambi-phones and hetero-phones, respectively [42].

Before we proceed, it is of interest to investigate the frequency with which different sounds are employed in Cantonese. This statistical study is based on the list of monosyllabic characters published by LSHK, together with supplementary characters provided by the Government of Hong Kong Special Administrative Region [39]. There are a total of 13,228 Cantonese pronunciations [43] of about 12,600 Chinese characters. Initially, we count the occurrence of each initial, F_i, final, F_f, and tone, F_t, in percentage. It is remarkable that 6 of these 19 initials, /l/, /g/, /h/, /z/, /c/, /s/ and /j/, are more popular, making up about 58% (58.3030%) of all Cantonese pronunciations. There are 196 characters (1.5528%) that start with a null initial (/0/). On the other hand, there are only 99 Chinese characters (0.7843%) that start with the initial, /kw/, which is the least popular initial. Table 2.1 displays a statistical analysis of the different initials used in Cantonese.

The “final” sounds are more evenly-distributed than the “initial” sounds for Cantonese. See Table 2.2 for the statistics. The more popular finals are /an/, /au/, /ai/, which make up 4.3179%, 4.5318%, and 4.6110% of all final sounds, respectively. The sounds ending with /up/, /en/, /um/, /ep/, /em/, /oet/, /m/, and /ng/ are less popular, and are only limited to the pronunciation of 0, 0, 1, 1, 1, 1, 1 and 1 Chinese character, respectively. There have six tones for Cantonese that same character may carry with more than one tone pronounced. All the tones pronounced for each character were counted. Finally, Table 2.3 displays the percentages of occurrence of the six tones used in the sample. The statistical analysis for Mandarin is shown in Table 2.4 to 2.6 that the nature of language of Cantonese is nearly the same as Mandarin.

2.3.4 Determination of corpora of daily usage Cantonese

Dewey [44] conducted a comprehensive study counting the frequency of the sounds of 100,000 English words. He used a wide variety of reading materials (corpora) in his study. No single source of his corpora contributed more than 5% of the total words. His corpora include editorial press reports and news published in different newspapers that were circulated at different cities in the U.S.A. The corpora also covered fiction, short stories, drama, speeches, personal correspondence, business correspondence, advertising materials, scientific English, religious English, magazine articles, and so forth.

Based on this study of the frequency of sounds of English words, Egan [4] outlined a systematic approach to establishing a database of 1,200 monosyllabic words. The database was used to construct word lists satisfying the following principles as far as practicable. They are: (1) a monosyllabic structure, (2) equal average difficulty, (3) equal range of difficulty, (4) equal phonetic composition, (5) a composition representative of English speech, and (6) words in common usage. Adapting these guiding principles, Zhang [37] established 750 PB characters for Mandarin. These two PB word/character lists subsequently formed the bases for the subjective speech intelligibility tests of English and Chinese (Mandarin), respectively.

Determining the most appropriate corpora for standard Cantonese poses a considerable problem in the current study. This is because Cantonese is mainly a spoken language due to the fact that Chinese people make a sharp distinction between written and spoken languages. This is in stark contrast with countries in the west, where such Romance languages as French, Italian and Spanish and such Saxon languages as English and German have a strong tendency to unify both the written and spoken forms of their respective languages.

Most importantly, the written form of Cantonese is based on the standard syntax and grammar of Mandarin. In Guangdong province of mainland China, Cantonese is usually written with the same set of characters as Mandarin, although the characters represent words not actually used in Cantonese. On the other hand, different words are often used (and hence written with different characters) in Hong Kong. Strictly speaking, Cantonese refers to the language spoken in and around the cities of Canton, presently known as Guangzhou, the capital city of Guangdong. Guangzhou was regarded as the centre of the purest form of Cantonese in the past. However, Hong Kong has now been widely accepted as the *de facto* “language” centre of Cantonese because of the influence of the mass media and pop culture originating from the city. As a result, the corpora for the construction of lists of PB Characters will be selected from sources originating in Hong Kong.

Cantonese is a colloquial language that is full of slang and non-standard usage. People in Hong Kong sometimes use the written form of spoken Cantonese in an informal way because it is more reflective and expressive, and received more warmly by speakers of

Cantonese. Most people use standard Chinese characters for formal occasions especially in records of legal documents, in order to capture exactly what a witness has said as official languages uses in Hong Kong for documents. As a result, the number of articles and documents in public circulation that use “pure” spoken Cantonese is not large. There are also noticeable differences in the ways Cantonese is spoken in Hong Kong and in other places around the world. In addition, the language of youth is rapidly evolving with new slang, and trendy expressions are constantly being introduced into mainstream spoken Cantonese. There is a strong tendency to mix English words or phrases with Cantonese during normal conversation, especially among the younger generation in Hong Kong. Consequently, Cantonese spoken in classrooms is more likely to be formal and official like written documents for public circulation such as newspapers or government announcements, than ‘pure’ spoken Cantonese with meaningless pronunciation or colloquial. Still, the selection of material is highly depended on the purpose of testing if the language’s nature is similar as Cantonese.

We re-iterate that 20 word lists are usually used in subjective tests to evaluate the speech intelligibility of English. Each of these lists contains 50 mono-syllabic words. These words are not chosen randomly but are selected in accordance with their relative frequencies of occurrence in daily usage. The distribution of phonological units, i.e. vowels and consonants, should be nearly the same in all word lists. The mono-syllabic words should be of equal difficulty in each list. By using different PB word lists, we can minimize the possible bias of hearing sensitivity on certain words or the frequency spectrum of the subjects during the speech intelligibility tests. This precaution can lead to more reliable experimental results without the repeated use of an identical word list.

2.3.5 Analysis of corpora of daily usage Cantonese

First we need to establish a set of appropriate corpora in selecting the PB character lists for Cantonese. Since the test materials for assessing the speech intelligibility in a communications channel should preferably be phonetically balanced, the database should be carefully chosen such that the test materials reflect the daily usage of a language. Local newspapers or even government announcements are good sources of the corpora for the development of such a paradigm. They are comparatively more reliable than books or journals. In the present study, the chosen corpora are comprised of 508 articles of local and international news published in different newspapers of Hong Kong between the three-month period from October to December in 1997 and 309 scripts of local TV news broadcasted in 2004. There have over 842 thousands monosyllabic words or Chinese characters with Cantonese pronunciation from the chosen corpora. They reflect the daily usage of spoken Cantonese and are used to analyze their phonological distribution: frequencies of initials, finals, and tones. It is worth re-iterating that each Chinese character is analogous to a mono-syllabic word in English. However, the textbooks have not been use as one of the sources for the corpora. The contents of the textbooks are Chinese characters in poems and literatures which may not pronounce in nowadays spoken Cantonese. Thus, the selected materials are not including those textbooks uses in schools.

Based on the list of monosyllabic characters published by LSHK, together with supplementary characters provided by the Government of Hong Kong Special Administrative Region [39], there are a total of 13,228 Cantonese pronunciations [43] of about 12,600 Chinese characters. The Cantonese pronunciations of each Chinese character in the corpora are firstly identified for calculation of their frequencies of occurrence. Although there have about 2,600 different Chinese characters in the corpora compares with about 13,000 published by LSHK, these characters are the most frequently used in our daily life.

It is noticeable that some Chinese characters have different tones when is use together with other Chinese character for different meaning. Hence, understanding its sentences' meaning is important for selection of appropriate tone for polyphone Chinese characters. For those are ambi-phones or heter-phones characters, the "normal" pronunciation are chosen according to LSHK. According to same statistical analysis on corpus published by LSHK, the frequency of occurrence for each initial, final and tone are summarized, together with their relative frequency of occurrence. The analysis of these frequencies is essential as base for development of a database of phonetically balanced words.

From the results shown in Table 2.7 to Table 2.9, and Figure 2.2

to Figure 2.4, the frequency of occurrence of each initial, final and tone are calculated. With reference to each initial, final and tone, a number is labelled to each for graphical presentation purpose only and their corresponding numbers are indicated in the tables. Both null initial or null final are represented into “0” in numbering. The graphical distributions are then plotted as shown in the figures.

Table 2.7 and Figure 2.2 shown /j/ is the most frequently used initial, with 13.073%. There are two possible reasons: 1) this initial is used in many common characters such as “one” /jat1/, “people” or “human” /jan4/, “have” /jau5/, it occurs about 10,000 times; 2) this initial is ranked as 6-th of the most frequent initial that has many characters are under this initial in pronunciation. The initial /s/ and /z/ follow with 12.376% and 11.738% respectively. Together with /g/, /d/ and /c/ that are ranked 4-th, 5-th and 6-th on the frequency of occurrence in the corpora, these top 6 initials contributed over 60% (61.054%) of populations. The least occurrences of initials are /0/ and /kw/, with 0.447% and 0.144% respectively. For initial /kw/, it occurs only 1,212 times in over 842,000 Chinese characters in the corpora. Compared with Table 2.1 and Table 2.7, the frequency distributions of initial are similar in the corpora and those from LSHK. Both have over half in population for /j/, /s/, /z/, /g/ and /c/ in total and are ranked as the most popular. The less popular are /0/ and /kw/. Figure 2.5 shows the comparison of percentage distributions of initial in the corpora and from LSHK.

Results from Table 2.8 and Figure 2.3 show that /i/ is the most popular final among 63 of them, with 7.715%. The main reason is that this final is included in common characters like “yes” or “is” /si6/ (5281 times), “market” /si5/ (5100 times), “time” or “period” /si4/ (3874 times) etc. The final /ing/ and /au/ are the occurrences in 2-nd and 3-rd in the corpora, with 5.570% and 5.101% respectively. These three most frequent finals cover 18.387% of distribution. Over 50% of contribution for the first 12 frequent finals and over 75% for the first 21 popular finals. Two-third of the finals occupied no more than 25% of distribution in the corpora. The least popular final is /ot/ with only 0.007%, however, /em/, /en/, /ep/, /et/, /eu/, /m/, /oe/, /oei/, /um/ and /up/ never occurs in the corpora. In comparison between Table 2.8 and Table 2.2, their similarity in frequency distribution is agreed and shown in Figure 2.6.

Table 2.9 shows that the most frequent tone is tone 1, with 25.722% and the 2-nd is tone 6, with 23.821%. These 2 tones contribute to about half of the population, 49.543%. The frequency distribution of tone in the corpora is similar as published by LSHK, which is shown in Figure 2.7. Analysis have been conducted in comparison to frequency distribution of initial, final and tone, all the results shown in Figure 2.5, Figure 2.6 and Figure 2.7 have positive regression. By comparing two distribution sources, their relative coefficient in similarity is 0.902 for initials. Under the deviation analysis, their coefficient is 0. Finals and tones are studied with the same conclusions. These results indicate there is no significant difference of distribution. With the above similarity characteristics, it implied that the selection of resources of the corpora is appropriated and, the corpora and those published by LSHK are assumed as unity resources. Therefore, the corpora used are appropriate for selection of phonetically balanced monosyllabic vocabulary in Cantonese and to reflect the daily usage of it.

The relative frequencies of occurrence of Chinese characters with Cantonese pronunciation have been compared with the respective results of Dewey [44] for English and Zhang [37] for Mandarin. This comparison is needed to minimize the bias in the process of selecting testing materials from the sources of corpora since wrong selection of the sources of corpora will lead to large deviation as bias on usage of certain characters (or syllables). First of all, the set of characters/words is ranked according to the relative frequencies of occurrence, F_n in percentage, of the characters/word. The character that occurs most frequently is ranked first, by setting $n = 1$. The next most frequently occurring character ranks second with $n = 2$, and so on. Figure 2.1 shows plots of the logarithmic relative frequencies of occurrence versus the logarithmic ranks for English, Mandarin, and Cantonese, respectively [37]. All three languages show different distributions in the relative frequencies although Cantonese and Mandarin show a rather similar trend except the extremely small or large occurrence since both languages come from the same language family. We expect these distributions will be dissimilar for different language families. The priority in selecting PB characters starts from the most frequently occurring character, i.e. $n = 1, 2, 3, . .$ and so on until all character lists have been completed.

Figure 2.1 shows the relative frequency of occurrence in English [44], Mandarin [37] and Cantonese. Since both Mandarin and Cantonese are tonal languages which come from same language family, it is clearly that Mandarin and Cantonese have comparable relationship in relative frequency and order of occurrences. In contrast, deviation between English and Cantonese is obvious, especially in range of higher orders and lower orders.

Selected characters are compiled in a list. Several lists of characters are established in the selection process. A further statistical analysis is conducted on each character list to determine the frequency of occurrence of each initial, final and tone. This serves to confirm all character lists are phonetically balanced and equally difficult. According to the results of distribution analyzed, frequency of occurrence of each initial, final and tone is allocated in order of relative occurrence n, started with n=1, 2, 3, .. and the simplest and the most commonly used Chinese characters are selected. The difficulties of the lists can be maintained and balanced.

According to the results of distribution analyzed, ten lists with 75 Chinese characters in Cantonese pronunciation have been chosen and shown in Appendix with equal difficulties and phonetically balanced. The number of Chinese characters in each list is the same as the one for Mandarin [30] also known as both languages come from same language family, polysyllabic and tonal language. In Hong Kong, and many cities using both spoken Chinese (Mandarin) and Cantonese in their societies, it is remarkable to use the lists for comparison of these two spoken languages for further study. In addition, 75 Chinese characters in each list can fulfil the distribution of percentages of initial, final and tone which 75 in total. Since these 10 lists should be phonetically balanced and same in frequency distribution for initial, final and tone, it is verified by statistical analysis on each lists and shown in Table 2.10 to 2.12. The characters will then be tested with a group of native speakers of Cantonese to ensure their appropriateness of the selected lists.

2.4 Conclusions

A set of 10 lists of phonetically balanced Chinese characters (with Cantonese pronunciation) has been chosen for subjective tests of speech intelligibility. There are 75 characters in each list and 10 lists have been chosen. The database can be used for more investigations with the aim of improving the speech intelligibility in classrooms and/or other communications channels. It aims to benefit the society for various purposes such as medical assessment on hearing systems, learning disabilities level, as monitor criteria for hearing impairment, and assessment of different kind of disabilities. It can further help on development of hearing aid equipment like audiphones. In this paper, the general methodology on the development of such a database has been discussed in detail. The methodology can be easily adapted by researchers or educators in other countries with different languages. The present study can be regarded as a guideline for developing pertinent information on phonetically balanced words for different languages.

References

1. Ethnologue: Languages of the World, 14th Edition. Ed., SIL International 2001 – 2005.
2. The United Nations Educational, Scientific and Cultural Organization (UNESCO) promotes linguistic diversity and multilingual education that aims to strengthen the unity and cohesion of societies. Starting from the millennium, UNESCO has been organizing an annual event known as International Mother Language Day.
3. EGAN, J. P., “Articulation Testing Methods,” Laryngoscope **58**, pp. 955-991, 1948.
4. FAIRBANKS, G., “Test of Phonemic Differentiation: The Rhyme Test,” J. Acoust. Soc. Am. **30**, pp. 596-600, 1958.
5. HOUSE, A. S., WILLIAMS, C. E., HECKER, M. H. L. and KRYTER, K. D., “Articulation Testing Methods: Consonantal Differentiation with a Closed-Response Set,” J. Acoust. Soc. Am. **37**, pp. 158-166, 1965.
6. WILLIAMS, C. E. and HECKER, M. H. L., “Relation between Intelligibility Scores for Four Test Method and Three Types of

Speech Distortion," J. Acoust. Soc. Am. **44**, pp. 1002 – 1006, 1968.

7. KEYTER, K. D. and WHITMAN, E. C., "Some Comparisons between Rhyme and PB-Word Intelligibility Tests," J. Acoust. Soc. Am. **37**, pp. 1146(L), 1965.
8. HOUTGAST, T. and STEENEKEN, H. J. M., "Evaluation of speech transmission channels by using artificial signals," Acustica **25**, pp. 355-367, 1971.
9. FRENCH, N. R. and STEINBURG, J. C., "Factors governing the intelligibility of sounds," J. Acoust. Soc. Am. **19**, pp. 90-119, 1947.
10. STEENEKEN, H. J. M. and HOUTGAST, T., "A physical method for measuring speed-transmission quality," J. Acoust. Soc. Am. **67**, pp. 318-326, 1980.
11. ANDERSON, B. W. and KALB, J. T., "English verification of the STI method for estimating speech intelligibility of a communications Channel," J. Acoust. Soc. Am. **81**, pp. 1982 – 1985, 1987.

12. KANG, J., "Comparison of speech intelligibility between English and Chinese," J. Acoust. Soc. Am., **103**, pp. 1213 – 1216, 1998.

13. Technical Notes and Briefs, "Device for measuring the speech transmission index," J. Acoust. Soc. Am. **71**, pp. 1612, 1982.

14. HOUTGAST, T. and STEENEKEN, H. J. M., "A multi-language evaluation of the RASTI-method for estimating speech intelligibility in auditoria," Acustica **54**, pp. 185 – 199, 1984.

15. SCHMIDT-NIELSEN, A., "Comments on the use of physical measures to assess speech intelligibility," J. Acoust. Soc. Am. **81**, pp. 1985 – 1987, 1987.

16. KOOTWIJK, P. A. A., "The speech intelligibility of the public address systems at 14 Dutch railway stations," J. Sound Vib. **193**, pp. 433 – 434, 1996.

17. IEC Standard 60268-16, 2nd Edition. Sound System equipment – Part 16: Objective rating of speech intelligibility by speech transmission index, 1998.

18. STEENEKEN, H. J. M., and HOUTGAST, T., "Mutual dependence of the octave band weights in predicting speech intelligibility," Speech Comm. **28**, pp. 109 – 123, 1999.

19. STEENEKEN, H. J. M., and HOUTGAST, T., “Phoneme-group specific octave-bands weights in predicting speech intelligibility,” Speech Comm. **38**, pp. 399 – 411, 2002.

20. GOVER, B. N. and BRADLEY, J. S., “Measures for assessing architectural speech security (privacy) of closed offices and meeting rooms,” J. Acoust. Soc. Am. **116**, pp. 3480 – 3490, 2004.

21. BENT, T. and BRADLOW, A. R., “The interlanguage speech intelligibility benefit,” J. Acoust. Soc. Am. **114**, pp. 1600 – 1610, 2003.

22. VAN WIJNGAARDEN, S. J., STEENEKEN, H. J. M., and HOUTGAST, T., “Quantifying the intelligibility of speech in noise for non-native talkers,” J. Acoust. Soc. Am. **112**, pp. 3004 – 3013, 2002.

23. VAN WIJNGAARDEN, S. J., STEENEKEN, H. J. M., and HOUTGAST, T., “Quantifying the intelligibility of speech in noise for non-native listeners,” J. Acoust. Soc. Am. **111**, pp. 1906 – 1916, 2002.

24. SATO, H., “Subjective measures to evaluate speech intelligibility, quality and difficulty in rooms for young and elderly listeners,” Can. Acoust. **30**, pp. 50 – 51 2002.

25. KISHON-RABIN, L. and ROSENHOUSE, J., “Speech perception test for Arabic speaking children,” Audiology **39**, pp. 269 – 277, 2000.

26. SKINNER, M. W., HOLDEN, L. K., HOLDEN, T. A., DEMOREST, M. E. and FOURAKIS, M. S, “Speech recognition at simulated soft, conversational, and raised-to-loud vocal efforts by adults with cochlear implants,” J. Acoust. Soc. Am. **101**, pp. 3766 – 3782, 1997.

27. MAGNUSSON, L., “Speech intelligibility index transfer functions and speech spectra for two Swedish speech recognition tests,” Scandinavian Audiology, **25**, pp. 59 – 67, 1996.

28. CLARK, J. E., “Four PB word lists for Australian English,” The Aust. J. Audiology **3**, pp. 21 – 31, 1981.

29. ANSI S3.2, American National Standard: Method for measuring the intelligibility of speech over communication systems, American National Standards Institute, New York, 1989.

30. GB/T 15508, National Standard of Peoples Republic of China: Acoustics – Speech articulation testing method, National Technical Supervision Bureau, 1995.

31. AS 2822, Australian Standard: Acoustics – Methods of assessing and predicting speech privacy and speech intelligibility, Standards Association of Australia, 1985.

32. Most linguists treat Cantonese as a separate language from Mandarin because of the different features that render these two languages are often mutually unintelligible. However, most Chinese speakers perceive Chinese to be one language partly because of the common written languages. Due to this self-perception, some linguists respect this terminology and use the word "dialect" for Cantonese. More details about the argument for treating Chinese as a family of languages can be found in "RAMEY, S. R. (1987), The Languages of China, Princeton University Press."

33. CHING, P. C., CHOW, K. F., LEE, T., NG, A. Y. P. and CHAN, L. W., "Development of a large vocabulary speech database for Cantonese," Proceedings of 1997 International Conference on

Acoustics, Speech and Signal Processing, Munich, Germany, April 1997.

34. LEE, T., CHING, P. C., CHAN, L. W., MAK, B. and CHENG, Y. H., "Tone recognition of isolated Cantonese syllables," IEEE Trans. on Speech and Audio Processing, Vol.3, No.3, pp. 204 – 209, May 1995.

35. LO, W. K., LEE, T. and CHING, P. C., "Development of Cantonese spoken language corpora for speech applications," Proceedings of 1998 International Symposium on Chinese Spoken Language Processing, pp. 102 – 7, Singapore, December 1998.

36. QIAN, Y., LEE, T. and SOONG, F. K., "Use of tone information in continuous Cantonese speech recognition," Proceedings of Speech Prosody 2004, pp. 587 – 590, Nara, Japan, March 2004.

37. ZHANG, J. L., edited by MAA, D. Y. and SHEN, H., Handbook of Acoustics, Chap. 19, pp. 404 – 435, Science Press, Beijing, 1987.

38. DEFRANCIS, J., The Chinese Language: Fact and Fantasy, Chap. 11 (The Monosyllabic myth), pp. 177 – 188, University of Hawaii Press, 1984.

39. The Linguistic Society of Hong Kong, Guide to LSHK Cantonese Romanization of Chinese Characters, The Linguistic Society of Hong Kong, 2002.
40. International Phonetic Association, Handbook of the International Phonetic Association: a guide to the use of International Phonetic Alphabet, Chap. 2, pp. 58 – 60, Cambridge University Press, Cambridge, 1999.
41. KANG, J., "Comparison of speech intelligibility between English and Chinese," J. Acoust. Soc. Am., **103**, pp. 1213 – 1216, 1998.
42. Chinese character database with word-formation phonologically disambiguated according to Cantonese dialect, The Chinese University of Hong Kong Research Center for Humanities Computing (see the website: http://www.arts.cuhk.edu.hk/Lexis/lexi-can/). As mentioned earlier, there are only about 1,700 separate tonal syllables.
43. DEWEY, G., Relative Frequency of English Speech Sounds, Harvard University Press, 1923.

Tables

Initial	/b/	/p/	/m/	/f/	/d/	/t/	/n/
Fi	532	429	563	547	592	543	257
Fi（%）	4.2149%	3.3988%	4.4605%	4.3337%	4.6902%	4.3020%	2.0361%

Initial	/l/	/g/	/k/	/ng/	/h/	/gw/	/kw/
Fi	883	932	443	287	817	223	99
Fi（%）	6.9957%	7.3839%	3.5097%	2.2738%	6.4728%	1.7668%	0.7843%

Initial	/w/	/z/	/c/	/s/	/j/	0[#]	Total
Fi	552	1327	998	1101	1301	196	12622
Fi（%）	4.3733%	10.5134%	7.9068%	8.7229%	10.3074%	1.5528%	100%

[#] null initials are represented 0.

Table 2.1 – The percentage distributions of the 20 initials (null initials included) of all Chinese characters with Cantonese pronunciations listed in Ref. [39].

Final	/i/	/ip/	/it/	/ik/	/im/	/in/	/ing/
Ff	510	111	201	332	224	424	500
Ff（%）	4.0406%	0.8794%	1.5925%	2.6303%	1.7747%	3.3592%	3.9613%

Final	/iu/	/yu/	/yut/	/yun/	/u/	/up/	/ut/
Ff	424	267	118	364	218	0	48
Ff（%）	3.3592%	2.1154%	0.9349%	2.8839%	1.7271%	0.0000%	0.3803%

Final	/uk/	/um/	/un/	/ung/	/ui/	/e/	/ep/
Ff	309	1	154	467	134	97	1
Ff（%）	2.4481%	0.0079%	1.2201%	3.6999%	1.0616%	0.7685%	0.0079%

Final	/et/	/ek/	/em/	/en/	/eng/	/ei/	/eu/
Ff	3	33	1	0	46	483	2
Ff（%）	0.0238%	0.2614%	0.0079%	0.0000%	0.3644%	3.8267%	0.0158%

Final	/eot/	/eon/	/eoi/	/oe/	/oet/	/oei/	/oek
Ff	53	189	452	5	1	2	94
Ff（%）	0.4199%	1.4974%	3.5810%	0.0396%	0.0079%	0.0158%	0.7447%

Final	/oeng/	/o/	/ot/	/ok/	/on/	/ong/	/oi/
Ff	275	316	17	240	87	440	177
Ff（%）	2.1787%	2.5036%	0.1347%	1.9014%	0.6893%	3.4860%	1.4023%

Final	/ou/	/ap/	/at/	/ak/	/am/	/an/	/ang/
Ff	508	95	249	67	265	545	164
Ff（%）	4.0247%	0.7527%	1.9727%	0.5308%	2.0995%	4.3179%	1.2993%

(cont’d)

Final	/ai/	/au/	/aa/	/aap/	/aat/	/aak/	/aam/
Ff	582	572	346	129	99	128	192
Ff（%）	4.6110%	4.5318%	2.7412%	1.0220%	0.7843%	1.0141%	1.5212%

Final	/aan/	/aang/	/aai/	/aau/	/m/	/ng/	0[#]	Total
Ff	330	86	222	187	1	1	34	12622
Ff（%）	2.6145%	0.6814%	1.7588%	1.4815%	0.0079%	0.0079%	0.2694%	100%

[#] null initials are represented 0.

Table 2.2 – The percentage distributions of the 63 finals (null finals included) of all Chinese characters with Cantonese pronunciations listed in Ref. [39].

Tone	1	2	3	4	5	6	Total
Ft	3297	1419	2199	2582	733	2392	12622
Ft（%）	26.121%	11.242%	17.422%	20.456%	5.807%	18.951%	100.000%

Table 2.3 – The percentage distributions of the 6 tones of all Chinese characters with Cantonese pronunciations listed in Ref. [39].

Initial	/b/	/p/	/m/	/f/	/d/	/t/	/n/	/l/	/g/	/k/	/h/	/j/	/q/
Fi（%）	5.15	0.98	3.74	2.45	12.00	3.53	2.53	5.69	5.50	1.83	4.42	6.98	3.11

Initial	/x/	/zh/	/ch/	/sh/	/r/	/z/	/c/	/s/	/o/
Fi（%）	4.86	7.18	2.75	7.66	1.94	3.01	1.15	1.08	12.45

Table 2.4 – The percentage distributions of the 22 initials of all Chinese characters with Mandarin pronunciations listed in Ref. [37].

Final	/a/	/o/	/e/	/i/	/u/	/ü/	/-i/	/er/	/ê/	/ai/	/ei/	/ao/	/ou/
Ff（%）	3.89	0.54	12.38	8.80	7.11	1.80	6.41	0.28	0.00	2.83	1.28	3.10	1.88

Final	/ia/	/ie/	/iao/	/iu/	/ua/	/uo/	/uai/	/ui/	/üe/	/an/	/en/	/ang/	/eng/
Ff（%）	1.09	2.42	2.06	2.60	0.44	4.40	0.32	2.75	1.01	3.41	3.62	2.87	3.09

Final	/ong/	/ian/	/in/	/iang/	/ing/	/iong/	/uan/	/un/	/uang	/ueng/	/üan/	/ün/
Ff（%）	4.18	4.10	1.95	1.80	3.05	0.42	1.24	0.89	0.65	0.003	0.85	0.52

Table 2.5 – The percentage distributions of the 38 finals of all Chinese characters with Mandarin pronunciations listed in Ref. [37].

Tone	1	2	3	4	5(null)
Ft（%）	18.71	19.37	17.51	35.79	8.63

Table 2.6 – The percentage distributions of the 5 tones of all Chinese characters with Mandarin pronunciations listed in Ref. [37].

Initial	/b/	/p/	/m/	/f/	/d/	/t/	/n/
Fi	41999	9360	35974	34335	73467	26526	13966
Fi（%）	4.9853	1.111	4.2702	4.0756	8.7206	3.1487	1.6578
X-axis	1	2	3	4	5	6	7

Initial	/l/	/g/	/k/	/ng/	/h/	/gw/	/kw/
Fi	38210	80597	19221	9046	44275	17466	1212
Fi（%）	4.5356	9.567	2.2816	1.0738	5.2555	2.0732	0.1439
X-axis	8	9	10	11	12	13	14

Initial	/w/	/z/	/c/	/s/	/j/	0[#]	Total
Fi	32750	98888	46999	104265	110132	3763	842451
Fi（%）	3.8875	11.738	5.5788	12.376	13.073	0.4467	100
X-axis	15	16	17	18	19	0	

[#] null initials are represented 0.

Table 2.7 – The percentage distributions of the 20 initials (null initials included) in the corpora.

Final	/i/	/ip/	/it/	/ik/	/im/	/in/	/ing/
Ff	64996	5276	6579	34270	4475	29933	46925
Ff（%）	7.7151	0.6263	0.7809	4.0679	0.5312	3.5531	5.5701
X-axis	1	2	3	4	5	6	7

Final	/iu/	/yu/	/yut/	/yun/	/u/	/up/	/ut/
Ff	17450	14806	5729	20902	10111	0	2126
Ff（%）	2.0713	1.7575	0.68	2.4811	1.2002	0	0.2524
X-axis	8	9	10	11	12	13	14

Final	/uk/	/um/	/un/	/ung/	/ui/	/e/	/ep/
Ff	12145	0	9866	31767	11653	7704	0
Ff（%）	1.4416	0	1.1711	3.7708	1.3832	0.9145	0
X-axis	15	16	17	18	19	20	21

Final	/et/	/ek/	/em/	/en/	/eng/	/ei/	/eu/
Ff	0	1095	0	0	829	38854	0
Ff（%）	0	0.13	0	0	0.0984	4.612	0
X-axis	22	23	24	25	26	27	28

Final	/eot/	/eon/	/eoi/	/oe/	/oet/	/oei/	/oek
Ff	5147	6457	21000	0	0	0	3868
Ff（%）	0.611	0.7665	2.4927	0	0	0	0.4591
X-axis	29	30	31	32	33	34	35

Final	/oeng/	/o/	/ot/	/ok/	/on/	/ong/	/oi/
Ff	28599	26015	57	17044	3886	22075	25155
Ff（%）	3.3947	3.088	0.0068	2.0231	0.4613	2.6203	2.9859
X-axis	36	37	38	39	40	41	42

(cont’d)

Final	/ou/	/ap/	/at/	/ak/	/am/	/an/	/ang/
Ff	39067	12302	23836	8110	9686	34226	9988
Ff（%）	4.6373	1.4603	2.8294	0.9627	1.1497	4.0627	1.1856
X-axis	43	44	45	46	47	48	49

Final	/ai/	/au/	/aa/	/aap/	/aat/	/aak/	/aam/
Ff	28670	42976	25056	2428	8883	8311	7596
Ff（%）	3.4032	5.1013	2.9742	0.2882	1.0544	0.9865	0.9017
X-axis	50	51	52	53	54	55	56

Final	/aan/	/aang/	/aai/	/aau/	/m/	/ng/	0[#]	Total
Ff	17109	3915	14768	6180	0	2550	0	842451
Ff（%）	2.0309	0.4647	1.753	0.7336	0	0.3027	0	100
X-axis	57	58	59	60	61	62	0	

[#] null initials are represented 0.

Table 2.8 – The percentage distributions of the 63 finals (null finals included) in the corpora.

Tone	1	2	3	4	5	6	Total
Ft	216697	92052	140483	137355	55184	200680	842451
Ft（%）	25.722	10.927	16.676	16.304	6.5504	23.821	100

Table 2.9 – The percentage distributions of the 6 tones in the corpora.

Initial	List 1	List 2	List 3	List 4	List 5	List 6	List 7	List 8	List 9	List 10
b	4	4	4	4	4	4	4	3	4	4
c	4	4	4	4	4	4	4	4	4	4
d	6	6	6	6	6	7	7	7	7	7
f	3	3	3	3	3	3	3	3	3	3
g	7	7	7	7	7	7	7	7	8	8
gw	2	2	2	2	2	2	1	1	1	1
h	4	4	4	4	4	4	4	4	4	4
j	10	10	10	10	10	10	10	10	10	10
k	2	2	2	2	2	2	2	1	1	1
kw	1	0	0	0	0	0	0	0	0	0
l	3	3	3	3	3	3	4	4	4	4
m	3	3	3	3	3	3	3	3	3	3
n	1	1	1	1	1	1	1	2	1	1
ng	1	1	1	1	1	1	0	0	1	1
p	1	1	1	1	1	1	1	1	0	0
s	9	9	9	9	9	9	9	10	10	10
t	2	2	2	2	2	2	3	3	3	3
w	3	3	3	3	3	3	3	3	3	3
z	9	9	9	9	9	9	9	9	8	8
0[#]	0	1	1	1	1	0	0	0	0	0
Total	75	75	75	75	75	75	75	75	75	75

[#] null initials are represented 0.

Table 2.10 – The percentage distribution of the 20 initials (null initials included) in 10 lists of phonetically balanced monosyllabic word lists in Cantonese.

Final	List 1	List 2	List 3	List 4	List 5	List 6	List 7	List 8	List 9	List 10
0[#]	0	0	0	0	0	0	0	0	0	0
aa	2	2	2	2	2	2	2	2	2	2
aai	1	1	1	1	1	1	2	2	2	2
aak	1	1	1	1	1	1	1	0	0	0
aam	1	1	1	1	1	1	1	0	0	0
aan	1	1	1	2	2	1	2	2	2	1
aang	0	0	0	0	0	0	0	1	1	1
aap	0	0	0	0	0	0	0	0	0	0
aat	1	1	1	1	1	1	1	1	0	0
aau	1	1	1	1	1	1	0	0	0	0
ai	3	3	3	3	3	2	2	2	2	2
ak	1	1	1	1	1	1	1	0	1	1
am	1	1	1	1	1	1	1	1	1	1
an	3	3	3	3	3	3	3	3	3	3
ang	1	1	1	1	1	1	1	1	1	1
ap	1	1	1	1	1	1	1	1	1	1
at	2	2	2	2	2	2	2	2	2	2
au	4	4	4	4	4	4	4	4	4	4
e	1	1	1	1	1	1	1	0	0	0
ei	4	4	4	3	3	3	4	4	3	3
ek	0	0	0	0	0	0	0	0	0	1
em	0	0	0	0	0	0	0	0	0	0
en	0	0	0	0	0	0	0	0	0	0
eng	0	0	0	0	0	0	0	0	0	1
eoi	2	2	2	2	2	2	2	2	2	2
eon	1	1	1	1	1	1	0	0	0	0
eot	0	0	0	0	0	1	1	1	1	1
ep	0	0	0	0	0	0	0	0	0	0
et	0	0	0	0	0	0	0	0	0	0
eu	0	0	0	0	0	0	0	0	0	0
i	6	6	6	6	6	6	6	6	6	6
ik	3	3	3	3	3	3	3	3	3	3
im	0	0	0	0	0	0	1	1	1	1
in	3	3	3	3	3	3	3	2	2	2
ing	4	4	4	4	4	4	4	4	4	4
ip	1	1	1	1	1	0	0	0	0	0

Final	List 1	List 2	List 3	List 4	List 5	List 6	List 7	List 8	List 9	List 10
it	1	1	1	1	1	1	0	0	0	0
iu	2	2	2	2	2	2	1	1	1	1
m	0	0	0	0	0	0	0	0	0	0
ng	0	0	0	0	0	0	0	0	1	1
o	2	2	2	2	2	2	2	3	3	3
oe	0	0	0	0	0	0	0	0	0	0
oei	0	0	0	0	0	0	0	0	0	0
oek	0	0	0	0	0	0	0	1	1	1
oeng	2	2	2	2	2	3	3	3	3	3
oet	0	0	0	0	0	0	0	0	0	0
oi	2	2	2	2	2	2	2	2	2	2
ok	2	2	2	2	2	1	1	1	1	1
on	0	0	0	0	0	0	0	1	1	1
ong	2	2	2	2	2	2	2	2	2	2
ot	0	0	0	0	0	0	0	0	0	0
ou	3	3	3	3	3	4	4	4	4	4
u	1	1	1	1	1	1	1	1	1	1
ui	1	1	1	1	1	1	1	1	1	1
uk	1	1	1	1	1	1	1	1	1	1
um	0	0	0	0	0	0	0	0	0	0
un	1	1	1	1	1	1	1	1	1	0
ung	3	3	3	3	3	3	3	3	2	2
up	0	0	0	0	0	0	0	0	0	0
ut	0	0	0	0	0	0	0	0	1	1
yu	1	1	1	1	1	1	1	2	2	2
yun	2	2	2	2	2	2	2	2	2	2
yut	0	0	0	0	0	1	1	1	1	1
total	75	75	75	75	75	75	75	75	75	75

[#] null initials are represented 0.

Table 2.11 – The percentage distribution of the 63 finals (null finals included) in 10 lists of phonetically balanced monosyllabic word lists in Cantonese.

Tone	List 1	List 2	List 3	List 4	List 5	List 6	List 7	List 8	List 9	List 10
1	20	20	19	19	19	19	19	19	20	19
2	8	8	9	8	8	8	8	8	8	8
3	12	12	12	13	12	13	13	13	12	13
4	12	12	12	12	13	12	12	12	12	12
5	5	5	5	5	5	5	5	5	5	5
6	18	18	18	18	18	18	18	18	18	18
Total	75	75	75	75	75	75	75	75	75	75

Table 2.12 – The percentage distribution of the 6 tones in 10 lists of phonetically balanced monosyllabic word lists in Cantonese.

Figures

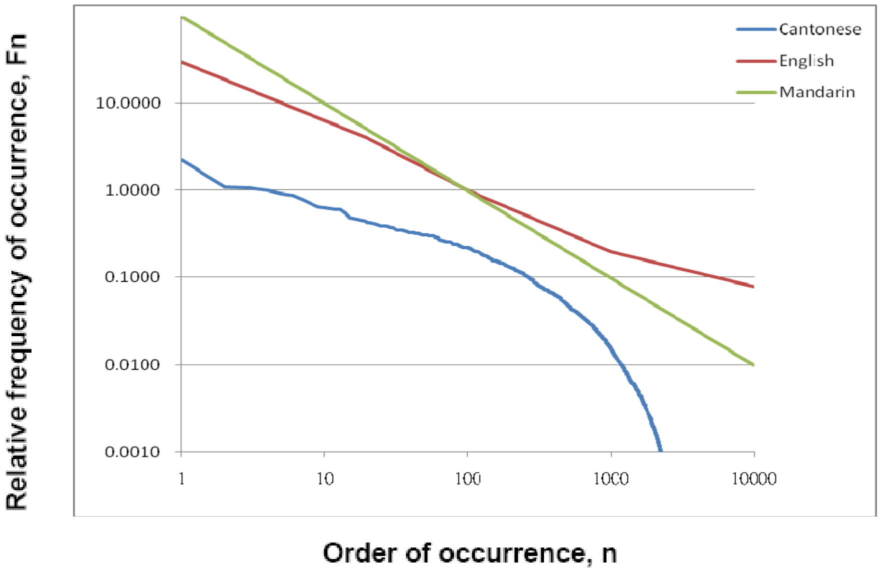


Figure 2.1 – Relative Frequencies of Occurrence with Ranking

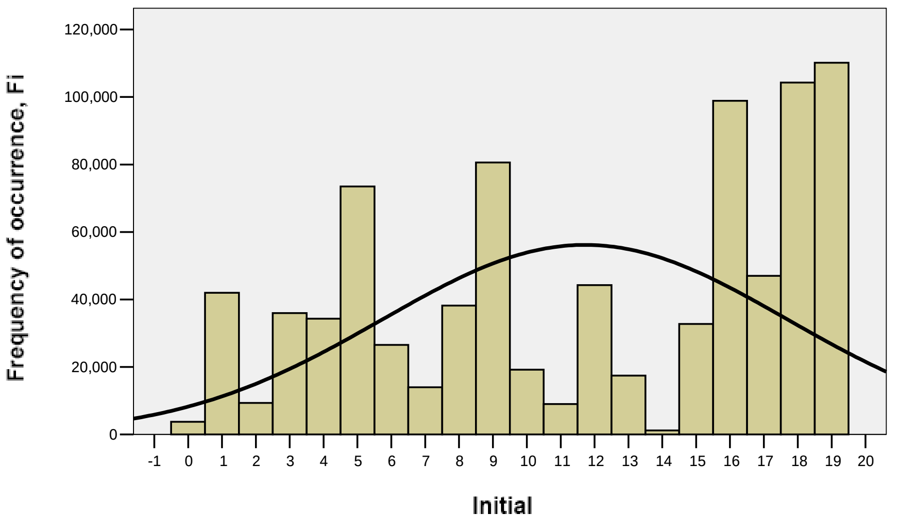


Figure 2.2 – The distributions of the 20 initials (null initials included) in the corpora.

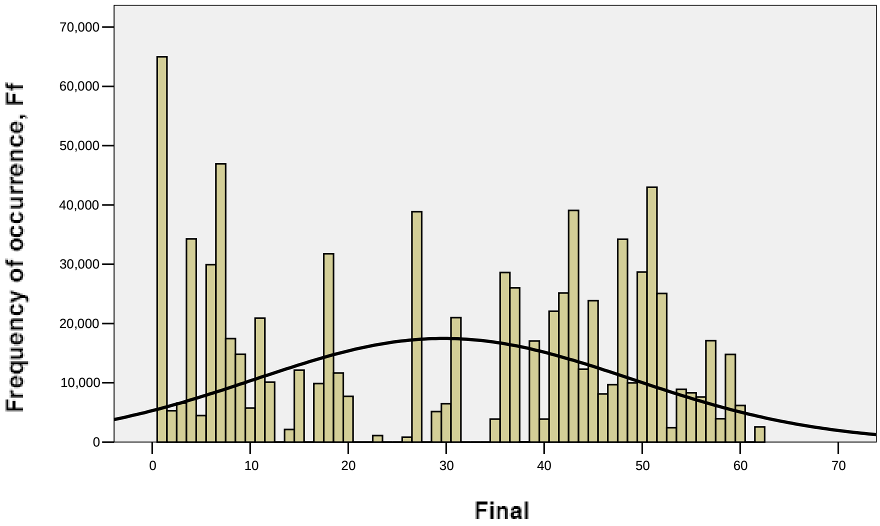


Figure 2.3 – The distributions of the 63 finals (null finals included) in the corpora.

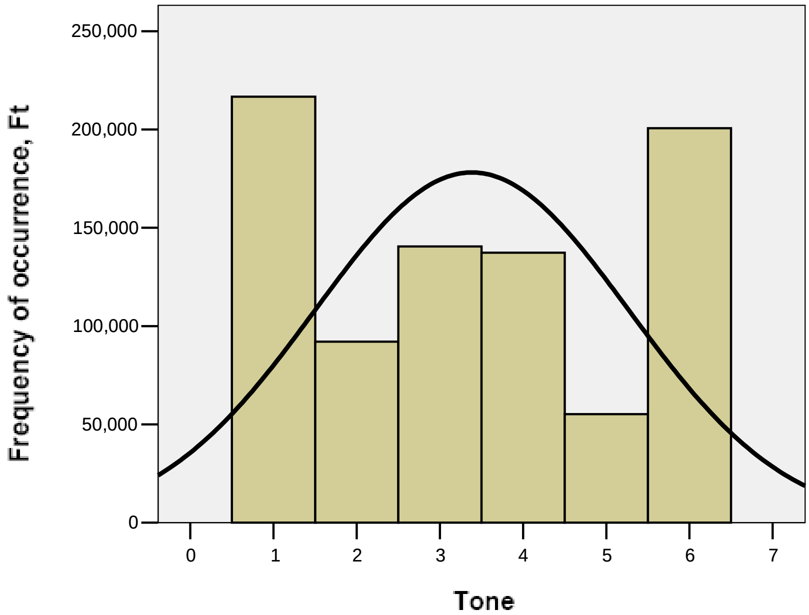


Figure 2.4 – The distributions of the 6 tones in the corpora.

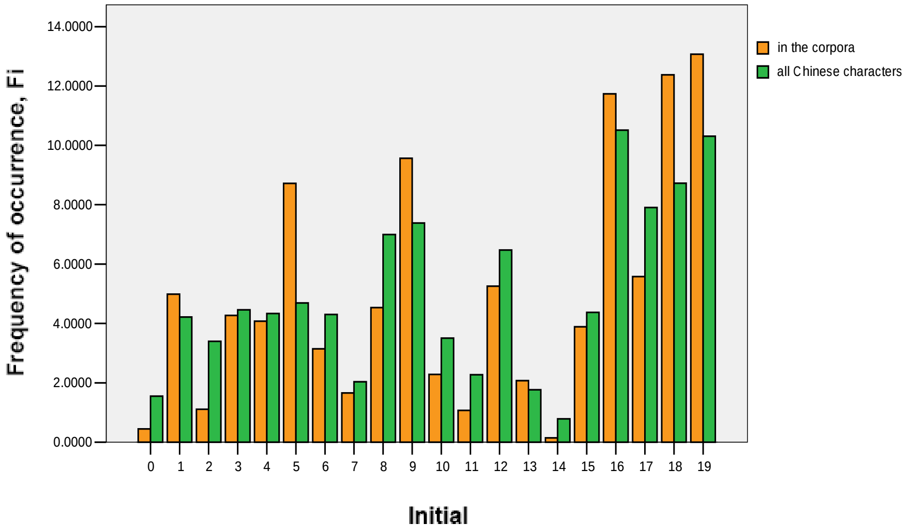


Figure 2.5 – The comparison of the distribution of 20 initials of the corpora and published by LSHK.

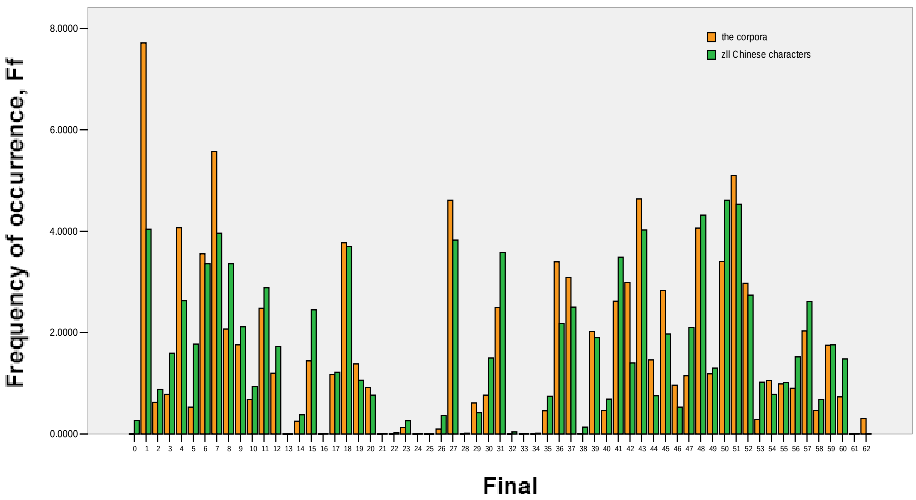


Figure 2.6 – The comparison of the distribution of 63 finals of the corpora and published by LSHK.

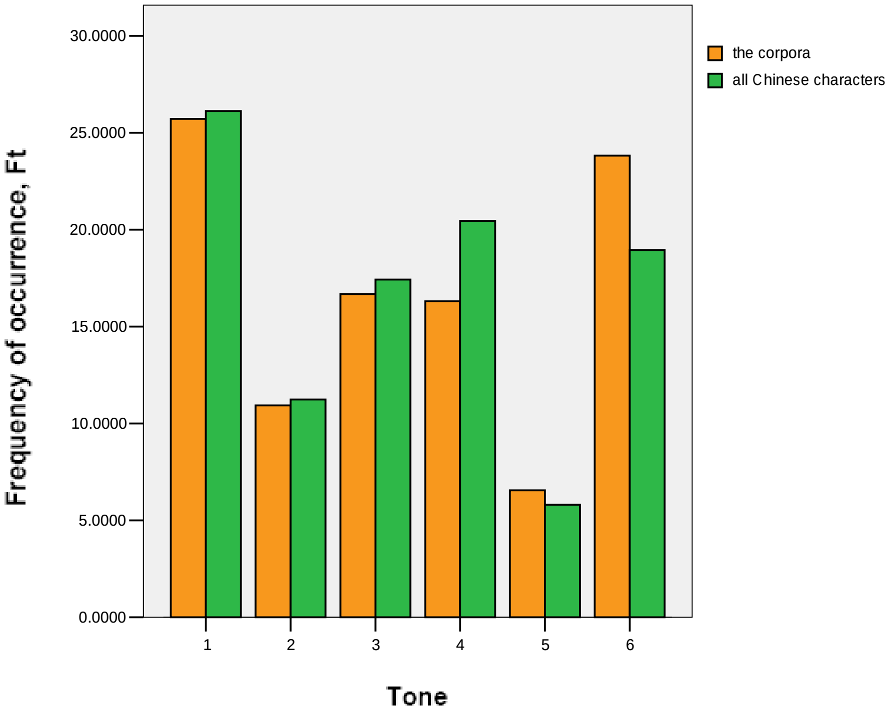


Figure 2.7 – The comparison of the distribution of 6 tones of the corpora and published by LSHK.

Appendix

List 1

百	baak3	堅	gin1	要	jiu1	是	si6
被	bei6	個	go3	於	jyu1	式	sik1
表	biu2	該	goi1	及	kap6	涉	sip3
本	bun2	港	gong2	強	koeng4	送	sung3
籌	cau4	高	gou1	擴	kwok3	他	taa1
晴	cing4	果	gwo2	離	lei4	體	tai2
場	coeng4	國	gwok3	里	lei5	爲	wai4
重	cung4	克	hak1	另	ling6	往	wong5
大	daai6	行	hang4	未	mei6	會	wui6
但	daan6	核	hat6	滅	mit6	洲	zau1
對	deoi3	券	hyun3	武	mou5	這	ze3
的	dik1	也	jaa5	女	neoi5	屯	zeon1
動	dung6	人	jan4	銀	ngan4	之	zi1
段	dyun6	一	jat1	片	pin3	子	zi2
發	faat3	又	jau6	三	saam1	政	zing3
否	fau2	兒	ji4	西	sai1	在	zoi6
府	fu2	益	jik1	晨	san4	造	zou6
交	gaau1	現	jin6	施	si1	續	zuk6
今	gam1	認	jing6	司	si1		

List 2

品	ban2	間	gaan1	用	jung6	市	si5
不	bat1	較	gaau3	預	jyu6	式	sik1
比	bei6	金	gam1	球	kau4	析	sik1
別	bit6	股	gu2	期	kei4	先	sin1
初	co1	官	gun1	拉	laai1	鮮	sin2
長	coeng4	季	gwai3	領	ling5	題	tai4
財	coi4	均	gwan1	料	liu6	頭	tau4
從	cung4	後	hau6	美	mei5	禍	wo6
得	dak1	器	hei3	明	ming4	匯	wui6
定	ding6	圈	hyun1	幕	mok6	濟	zai3
當	dong1	犬	hyun2	南	naam4	者	ze2
倒	dou2	有	jau5	能	nang4	最	zeoi3
毒	duk6	意	ji3	外	ngoi6	資	zi1
動	dung6	疑	ji4	澳	ou3	仔	zi2
化	faa3	義	ji6	普	pou2	接	zip3
法	faat3	亦	jik6	十	sap6	將	zoeng1
分	fan1	現	jin6	室	sat1	作	zok3
加	gaa1	營	jing4	需	seoi1	裝	zong1
格	gaak3	擾	jiu5	順	seon6		

List 3

包	baau1	計	gai3	勇	jung5	時	si4
品	ban2	吉	gat1	與	jyu5	事	si6
避	bei6	購	gau3	區	keoi1	升	sing1
並	bing6	基	gei1	權	kyun4	上	soeng6
車	ce1	關	gwaan1	力	lik6	天	tin1
除	ceoi4	軍	gwan1	連	lin4	條	tiu4
清	cing1	捐	gyun1	了	liu5	威	wai1
請	cing2	下	haa6	民	man4	和	wo4
打	daa2	客	haak3	亡	mong4	胡	wu4
達	daat6	厚	hau5	姆	mou5	增	zang1
特	dak6	響	hoeng2	尼	nei4	進	zeon3
地	dei6	入	jap6	樂	ngok6	質	zi3
跌	dit3	日	jat6	愛	oi3	自	zi6
多	do1	醫	ji1	舖	pou3	積	zik1
方	fong1	研	jin4	勢	sai3	直	zik6
悔	fui3	業	jip6	心	sam1	助	zo6
款	fun2	育	juk6	手	sau2	災	zoi1
界	gaai3	擁	jung2	受	sau6	早	zou2
減	gaam2	容	jung4	試	si3		

List 4

價	gaa3	七	cat1	宜	ji4	載	zoi3
炸	zaa3	筆	bat1	息	sik1	確	kok3
帶	daai3	就	zau6	職	zik1	惡	ok3
劃	waak6	救	gau3	年	nin4	望	mong6
參	caam1	偷	tau1	戰	zin3	廣	gwong2
萬	maan6	候	hau6	善	sin6	度	dou6
還	waan4	些	se1	成	sing4	組	zou2
八	baat3	利	lei6	精	zing1	導	dou6
教	gaau3	死	sei2	影	jing2	夫	fu1
提	tai4	技	gei6	評	ping4	恢	fui1
偉	wai5	里	lei5	葉	jip6	獨	duk6
幣	bai6	去	heoi3	列	lit6	門	mun4
北	bak1	隨	ceoi4	喬	kiu4	東	dung1
感	gam2	詢	seon1	邀	jiu1	融	jung4
新	san1	以	ji5	過	gwo3	風	fung1
很	han2	治	zi6	俄	ngo4	如	jyu4
根	gan1	使	si2	長	zoeng2	軟	jyun5
更	gang3	似	ci5	央	joeng1	院	jyun6
合	hap6	思	si1	代	doi6		

List 5

班	baan1	記	gei3	員	jyun4	示	si6
爆	baau3	居	geoi1	遠	jyun5	星	sing1
標	biu1	擊	gik1	求	kau4	承	sing4
播	bo3	沽	gu1	確	kok3	她	taa1
察	caat3	公	gung1	立	lap6	臺	toi4
測	cak1	季	gwai3	樓	lau4	屋	uk1
遲	ci4	光	gwong1	兩	loeng5	或	waak6
程	cing4	系	hai6	無	mou4	位	wai6
擔	daam1	起	hei2	每	mui5	雲	wan4
等	dang2	顯	hin2	們	mun4	陣	zan6
逗	dau6	協	hip6	我	ngo5	知	zi1
調	diu6	任	jam6	內	noi6	置	zi3
黨	dong2	伊	ji1	偏	pin1	績	zik1
到	dou3	爾	ji5	實	sat6	展	zin2
反	faan2	疫	jik6	收	sau1	節	zit3
費	fai3	仍	jing4	寫	se2	昨	zok6
佛	fat6	陽	joeng4	四	sei3	做	zou6
家	gaa1	用	jung6	水	seoi2	中	zung1
近	gan6	餘	jyu4	信	seon3		

List 6

巴	baa1	機	gei1	爾	ji5	食	sik6
辦	baan6	舉	geoi2	億	jik1	線	sin3
備	bei6	建	gin3	魚	jyu4	城	sing4
半	bun3	京	ging1	級	kap1	少	siu2
策	caak3	勁	ging6	抗	kong3	鐵	tit3
產	caan2	高	gou1	林	lam4	台	toi4
陳	can4	貴	gwai3	論	leon6	話	waa6
曾	cang4	國	gwok3	率	leot6	爲	wai6
島	dou2	效	haau6	物	mat6	和	wo4
度	dou6	何	ho4	面	min6	暫	zaam6
督	duk1	向	hoeng3	名	ming4	則	zak1
董	dung2	空	hung1	尼	nei4	州	zau1
短	dyun2	因	jan1	外	ngoi6	滋	zi1
段	dyun6	引	jan5	捧	pung2	至	zi3
奪	dyut6	一	jat1	殺	saat3	字	zi6
放	fong3	由	jau4	售	sau6	照	ziu3
夫	fu1	有	jau5	社	se5	張	zoeng1
恢	fui1	儀	ji4	雖	seoi1	遭	zou1
解	gaai2	議	ji5	適	sik1		

List 7

備	bei6	據	geoi3	用	jung6	識	sik1
部	bou6	警	ging2	雨	jyu5	消	siu1
貝	bui3	改	goi2	禽	kam4	所	so2
般	bun1	各	gok3	決	kyut3	彈	taan4
查	caa4	降	gong3	量	loeng6	圖	tou4
且	ce2	君	gwan1	路	lou6	團	tyun4
吹	ceoi1	卷	gyun2	六	luk6	挽	waan5
出	ceot1	後	hau6	龍	lung4	維	wai4
第	dai6	氣	hei3	買	maai5	混	wan6
德	dak1	汽	hei3	盟	mang4	責	zaak3
點	dim2	可	ho2	微	mei4	站	zaam6
電	din6	一	jat1	年	nin4	真	zan1
待	doi6	油	jau4	派	paai3	周	zau1
都	dou1	右	jau6	薩	saat3	指	zi2
冬	dung1	演	jin5	失	sat1	致	zi3
花	faa1	英	jing1	斯	si1	自	zi6
方	fong1	應	jing3	市	si5	值	zik6
符	fu4	樣	joeng6	士	si6	正	zing3
急	gap1	讓	joeng6	色	sik1		

List 8

便	bin6	蓋	goi3	越	jyut6	喪	song3
波	bo1	角	gok3	卡	kaa2	數	sou3
報	bou3	講	gong2	留	lau4	屬	suk6
層	cang4	告	gou3	理	lei5	書	syu1
取	ceoi2	沽	gu1	類	leoi6	太	taai3
持	ci4	骨	gwat1	零	ling4	同	tung4
赤	cik3	口	hau2	羅	lo4	灣	waan1
單	daan1	香	hoeng1	馬	maa5	穩	wan2
低	dai1	開	hoi1	問	man6	緩	wun6
突	dat6	看	hon3	媒	mui4	扎	zaat3
定	ding6	引	jan5	你	nei5	際	zai3
道	dou6	醫	ji1	寧	ning4	執	zap1
動	dung6	易	ji6	疲	pei4	就	zau6
短	dyun2	嚴	jim4	生	saang1	支	zi1
快	faai3	演	jin5	深	sam1	只	zi2
非	fei1	弱	joek6	術	seot6	席	zik6
科	fo1	容	jung4	師	si1	召	ziu6
九	gau2	遇	jyu6	性	sing3	像	zoeng6
激	gik1	元	jyun4	相	soeng1		

List 9

敗	baai6	繼	gai3	燕	jin3	商	soeng1
北	bak1	件	gin6	約	joek3	常	soeng4
步	bou6	經	ging1	月	jyut6	送	sung3
勃	but6	歌	go1	其	kei4	殊	syu4
親	can1	乾	gon1	蘭	laan4	坦	taan2
次	ci3	估	gu2	律	leot6	透	tau3
刺	cik3	局	guk6	淩	ling4	拖	to1
全	cyun4	怪	gwaai3	涼	loeng4	華	waa4
登	dang1	希	hei1	命	ming6	溫	wan1
隊	deoi6	海	hoi2	務	mou6	回	wui4
代	doi6	學	hok6	滿	mun5	爭	zaang1
當	dong1	紅	hung4	紐	nau2	制	zai3
刀	dou1	音	jam1	五	ng5	質	zat1
杜	dou6	一	jat1	什	sap6	周	zau1
斷	dyun6	友	jau5	歲	seoi3	就	zau6
飛	fei1	而	ji4	試	si3	之	zi1
火	fo2	以	ji5	釋	sik1	自	zi6
防	fong4	易	jik6	聲	sing1	主	zyu2
架	gaa3	染	jim5	小	siu2		

List 10

病	beng6	奇	gei1	約	joek3	述	seot6
保	bou2	檢	gim2	羊	joeng4	視	si6
布	bou3	競	ging6	娛	jyu4	想	soeng2
盃	bui1	哥	go1	曲	kuk1	殊	syu4
輯	cap1	港	gong2	冷	laang5	天	tin1
青	cing1	沽	gu1	流	lau4	停	ting4
唱	coeng3	功	gung1	歷	lik6	託	tok3
賽	coi3	棍	gwan3	來	loi4	懷	waai4
地	dei6	克	hak1	晚	maan5	域	wik6
兌	deoi6	可	ho2	文	man4	王	wong4
隊	deoi6	寒	hon4	密	mat6	債	zaai3
訂	ding3	好	hou3	午	ng5	製	zai3
到	dou3	也	jaa5	嫩	nyun6	增	zang1
凍	dung3	任	jam6	沙	saa1	走	zau2
端	dyun1	伊	ji1	西	sai1	紙	zi2
奪	dyut6	移	ji4	身	san1	織	zik1
乏	fat6	已	ji5	首	sau2	招	ziu1
非	fei1	二	ji6	受	sau6	坐	zo6
闊	fut3	言	jin4	石	sek6		

Chapter 3
Acoustical conditions of classrooms of selected schools in Hong Kong

3.1 Introduction

Hong Kong is currently in the midst of the largest campaign of stepping up the quality of school education in history. The government has launched a school improvement programme not only to re-provision the existing schools, also to revise the standard design for primary and secondary schools. This programme aims to provide more facilities to support the implementation of various new initiatives and enhancing language teaching. Further than that, it is not only the case that happens in Hong Kong but in all the rapidly growing countries in Asia such as China or Middle East. In fact, effective verbal communication between teachers and students in classrooms is very important and beneficial to both. An optimal learning environment is crucial to the development of characters and learning patterns of the students. The teachers are not required to increase their voice to guarantee the communication of instructions to the students. The usual symptoms like headache and sore throat of the teachers caused by compensating the poor acoustical conditions by raising their voice could also be minimised. The students' hearing system can be protected without improper usage of amplifiers or address systems inside the classrooms.

Comprehensive acoustical measurements of existing classrooms for different age groups from primary to tertiary educations in Hong Kong have been conducted. The analysed data provided valuable information on the empirical relationship between the speech intelligibility in a classroom and the variations in the reverberation time, the background noise level, the signal-to-noise ratio, and the speech transmission index. The outcome of these experimental results provides an insight on the acoustical conditions of the typical classrooms in Hong Kong that compared with literature review of the worldwide standards for the design criteria.

3.2 Acoustical measurements of classrooms in Hong Kong

To establish an effective learning environment, speech intelligibility in classrooms should be improved to achieve satisfactory acoustical conditions or up to international standards, at least. However, it is much more expensive and difficult in practice to renovate the built classrooms. The vital importance is to identify the acoustical design criteria for the classrooms beforehand. Even the issues of "good" acoustical conditions and "adequate" speech transmission inside classrooms have been discussed for years but little, if anything, have been investigated. Early studies investigated that speech intelligibility can be interpreted by indoor ambient noise level (or background noise level) and reverberation inside classrooms [1 – 3]. Excessive indoor noise and reverberation inside classrooms interfere with the speech intelligibility affecting the influence on behaviour and achievement of the students, and so lead to health problems of either the teachers or the students.

Several acoustical properties that affect verbal communication have been carried out by experimental measurements. These properties include outdoor and indoor ambient noise levels, noise criteria, and reverberation time in 500Hz, 1000Hz, 2000Hz and its arithmetic average. In addition to quantify the acoustical properties of the classrooms surveyed, the well-known objective speech metrics called speech transmission index [1 – 3] and its simplified version – rapid speech transmission index [4 – 5] have been taken into account in order to influence the speech intelligibility.

3.2.1 Background information of the classrooms

Twenty classrooms were selected from five primary schools, five secondary schools and a university for field measurements. The classrooms from primary and secondary schools were designed and built based on modern designs of the government in Hong Kong. Most of them were built under a standardized design named ‘generation 2000’ or ‘millennium generation’ (G2000) blueprint. Since the design is only applied for primary and secondary schools in Hong Kong, all universities were built on their own design. Though, such blueprint has extent to use in new generation of schools in China, for example, Shanghai. For comparison of their acoustical conditions, the classrooms selected in the university are comparatively similar in geometrical configuration with the classrooms in primary and secondary schools.

Table 3.1 lists out the classrooms measured with their architectural and geometrical details. Some of them are finished with acoustic treatments like carpet, acoustics ceiling or acoustics back wall panel which are highlighted in the table. The number of split type air-conditioning (A/C) machines and fresh air pre-conditioning (FAP) machines in each classroom are recorded for discussion purpose of noise criteria.

From the selected primary schools, there are five classrooms which were built with G2000 design and one of which was finished with acoustics ceiling. Four classrooms with non-standard design were built with acoustics back panel. In addition, all measured classrooms from selected secondary schools were built with G2000 design and two of them were finished with acoustics ceiling. Four measured university classrooms were built with non-standard design with highly similar geometrical configuration; however, all of them were paved with carpet.

3.2.2 Experimental procedures of field measurements

The acoustical measurements were deliberately conducted in school hours to reflect the influence of background noises, such as traffic noise and student activities, on the acoustical conditions of classrooms under normal operation conditions. The measurements were carried out under unoccupied condition by three diffuse field microphones and two free field microphones. The measurements were repeated to account for the reproducibility of results in different sets of data whenever possible. Figure 3.1 shown details of locations of the microphones. The diffuse field microphones numbered 1 to 3 were placed inside the classrooms diagonally to measure the acoustical parameters. Meanwhile, the free field microphones numbered 4 and 5 were located at corridor and outside the windows respectively to monitor the outdoor noises. All of the microphones inside classrooms were set at 1.2 m above the floor level which is nearly at ear level of the seated students. The noise signals received from the microphones were recorded simultaneously in a Sony 16-channel digital recorder for all sets of measurements. The signals recorded were downloaded and analyzed together with the 01dB-Stell Signal/Frequency Analyzer and Building Acoustics Analyzer which is used to calculate the acoustical parameters and the speech indices.

Two normal conditions of the classrooms have been operated to measure the indoor ambient noise level and noise criteria in 5 minutes duration. The normal conditions have two different ventilation operation systems. The first condition for measurement was carried out with the doors closed, the mechanical ventilation system was switched off but all the windows were widely open. This case is the most usual operation condition of the classroom in autumn or winter in Hong Kong. The second condition was conducted with all the windows and doors closed but all mechanical equipments for ventilation switched on. This is the most usual operation condition of the classroom in spring or summer in Hong Kong. Since the study was focused on improving the speech intelligibility of the classrooms, the real and practical situation should be used for investigation. These two conditions are almost the operation condition for whole academic year. The measured data from this case were used to calculate the noise criteria of the classroom in which the most affecting noise sources were the mechanical equipments for ventilation only.

The measurements of reverberation time (RT), speech transmission index (STI) and the rapid speech transmission index (RASTI) were conducted by the impulse which generated by pricking an inflated balloon. All windows were closed and the mechanical ventilation system (A/C & FAP) was switched off for these measurements. The inflated balloons were placed about 1.5 m above the floor, which is about the mouth level of an adult, at three positions for bursting. Three impulses were generated in front of blackboards and cater-corners of the classrooms where cross diagonal of the diffuse field microphones. Figure 3.1 shows the details of location of these three bursts. Each measurement lasted for about 12 seconds for sufficient decay of signal inside the classroom. The results of RT, STI and RASTI at each microphone's position were obtained by these three sets of balloon burst at the described location.

3.3 Comparison of international standards

In recent years, the classroom acoustics has aroused a great concern internationally and many countries have developed or revised standards for acoustical design of classrooms. We endeavour to identify the international standards, guidelines and recommendations of different countries and organizations for the design of classrooms in the study. These international standards are based on the studies initiated by Vallet [6 – 7] in eight European countries: Belgium [8], Germany [9 – 10], United Kingdom [11], France, Turkey, Sweden [12], Italy and Portugal. Five countries: USA [13], Canada [14], China [15], Australia and New Zealand [16] are included in this study to cover the studies on the region of Pacific Rims. The study also includes recommendation from the World Health Organization (WHO) [17], the American Society of Heating, Refrigeration and Air Conditioning Engineers (ASHARE) [18], and the American Speech Language Hearing Association (ASHA) [19].

The ultimate goal in classroom acoustics is the adequate speech intelligibility. This is known to be mainly determined by the signal-to-noise level difference (S/N): the difference between the speech-signal and background-noise levels at a receiver, and the amount of reverberation [1 – 3]. The usual way to characterize the amount of reverberation in a room is by the reverberation time (RT). Reverberation time and ambient noise level are therefore the usual basis of classroom acoustics standards and regulations. Despite the fact that most of the speech metrics, i.e. STI and RASTI, that have been proposed can be measured in real rooms, classroom acoustical performance is seldom specified in terms of these quantities. In these international standards, guidelines and recommendations, 3 usual parameters are identified for achieving optimal acoustical designs of classrooms. These 3 acoustical parameters are indoor noise levels, reverberation time, and noise criteria (NC). Most of these standards specify two parameters (ambient noise levels and reverberation time). However, ASHARE specifies NC only and ASHA specifies all three parameters. The recommended values of these three acoustical parameters for the optimal acoustic design of classroom are summarised in Table 3.2. The acoustical design criteria specified by France, Italy, Portugal and Turkey as shown in Table 3.2 are referred to Vallet's paper [6 – 7] since the original documents are not accessible by the authors.

In spite of the limits for indoor ambient noise level are various in different countries and organizations which are between 30 and 50 in A-weighted equivalent sound pressure levels, most of them specify the limit at 35 dB(A) whilst ASHA gives the most stringent requirement and China permits the greatest noise tolerance. In France, the noise limits are 33 dB(A) and 38 dB(A) when the noise is continuous and discontinuous respectively. Germany permits a 5 dB(A) higher in noise limit for HVAC equipment as long as this noise is continuous without distinctly audible single tones. The American National Standard Institute (ANSI) specifies 40 dB(A) for the noise limit when the nosiest hours is dominated by the transportation noise. With regard to the control of reverberation, most of the international standards and guidelines specify maximum reverberation times between 0.4 and 0.8 second in the octave bands with mid-band frequencies of 500Hz, 1000Hz or 2000Hz, or the arithmetic average of the reverberation time (T_{mf}) in these frequency bands with the exception of China where a greater reverberation is permissible. The reverberation times as specified by Belgium, United Kingdom, Australia and New Zealand, ASHA and ANSI are for unoccupied classrooms. However, it is not clear whether the requirements of reverberation time apply to unoccupied or occupied classrooms in the standards and guidelines of France, Portugal, Canada and WHO.

Among all the countries and organizations, only ASHA and ASHRAE restricts the equipment noise level by specifying NC rating at NC 20 and NC 30 respectively.

3.4 Results of field measurements

3.4.1 Indoor ambient noise level and outdoor noise level

Figure 3.2 (a) and 3.2 (b) show the measured results of the indoor ambient noise levels inside the classrooms and their instantaneous outdoor noise levels respectively. Both measured results were conducted in a 5 minutes measurement period with the windows open for ventilation and the windows closed in the air-conditioning environment. Since the overall indoor ambient noise level is the objective and essential parameter for design stage of a classroom, the presented results are averaged of measured data from position 1 to 3 inside the classroom. The outdoor noise levels were logarithmic average of the measured signals from position 4 and 5 where placed at corridor and outside the windows.

As shown in Figure 3.2 (a), all measured classrooms, except one classroom (P2 – A) from primary schools, exceeded 50 dB (A) in indoor ambient noise level under both operation conditions. By comparing the corresponding averaged outdoor noise level, it was relatively low in noise level for that exceptional classroom during the measurement which about 55 to 58 dB (A). Most measured classrooms have indoor ambient noise levels in the range of 50 to 60 dB (A) in both operation conditions while the averaged outdoor noise levels in the range of 65 to 75 dB (A). However, three of them (S1 – A, S1 – B, S3 – A) have the indoor ambient noise levels higher than 60 dB (A) under the condition of natural ventilation. Two classrooms (P2 – B, P4 – A) from primary schools and two classrooms (U1 – A, U1 – B) from university have nearly 50 dB (A) in indoor ambient noise level under condition of mechanical induced ventilation which all windows were closed for the air-condition environment.

Most of the measured classrooms have higher indoor noise levels under the operation condition of natural ventilation than the mechanical induced one. The apparently high indoor ambient noise levels were largely due to the large outdoor background noise from the measuring environments in the vicinity of classrooms. With all windows opened for natural ventilation, the indoor ambient noise levels depended on the outdoor noise levels at the moment. From the results of the classrooms P2 – B, S3 – A, U1 – C and U1 – D, the averaged outdoor noise levels are higher during the measurement period for air-conditioning environment than the natural ventilation. Nevertheless, the simultaneous indoor ambient noise levels under condition of air-conditioning environment are apparently lower than under windows opened ventilation. The function of windows closed to increase the transmission loss is obvious. The windows were used as sound barrier to insulate the outdoor background noise when they were closed so that the ambient noise levels inside the classrooms were apparently lower than when the windows were opened. The discussion could further be illustrated by comparing the results of condition of natural ventilation with air-conditioning environment and with the instantaneous averaged outdoor noise levels of the measured classrooms.

What is more, it is worth noting that in the classroom P2-A, there was an apparent increase in the outdoor noise levels due to the temporarily increased of traffic noise during the measurements. The indoor ambient noise level under condition of mechanical ventilation was higher compared with other measurements even when the classroom's windows were open for natural ventilation in the latter cases. The discussion on this data was discarded but was retained for completeness purpose.

Although the indoor ambient noise levels in all measured classrooms were near or below 60 dB (A) which seems an acceptable limit for most people, it is still unsatisfactory to meet any national or industrial standards or guidelines. It is worth noting that the indoor ambient noise level from the international standards or guidelines were 35 dB (A) at most. The quietest classroom (P2 – A) measured has reluctantly fulfilled grade B of the national standard from China. It is understood the acoustical condition of classroom is more stringent for learning efficiently through an 'adequate' speech transmission.

3.4.2 Noise criteria

The discussion above shows the air-conditioning environment is quieter than the natural ventilation inside the classrooms measured, still, there are another noise sources affecting the acoustical condition of the classrooms. Noise radiated from the mechanical equipment inside the classrooms inevitably deteriorates the speech transmission between teachers and students. Figure 3.3 shows the noise criteria (NC) of the classrooms under the operation condition with air-conditioning environment where all the windows were closed and the mechanical ventilation system was switched on without presence of the students. The figure shows the averaged values of the measured data from position 1 to 3 under a 5 minutes´ measurement period.

Most classrooms under testing have the noise criteria nearly or below NC 50. Only five classrooms (P1 – A, P1 – B, P5 – B, S1 – A, S4 – A) out of twenty have NC 55 and there is one classroom (S2 – A) with noise criteria nearly NC 60. The abnormally high of noise criteria of the classroom can attribute to the extraordinarily high averaged outdoor noise level during the measurement of corresponding condition. The difference on indoor ambient noise level between condition of natural ventilation and mechanical induced ventilation is about 8 dB (A) while there have 10 dB (A) difference on averaged outdoor noise level.

With reference to the organizations shown in Table 3.2, all the classrooms measured have the noise criteria exceed the recommended value, NC 30 even the quietest classroom (P2 – A) measured have NC 45. This may ascribed to comparatively high of indoor ambient noise levels of the classrooms measured with the international standards and organizations.

3.4.3 Reverberation time

The reverberation times at the frequencies of 500 Hz, 1000 Hz and 2000 Hz at different positions inside the classrooms are given in Figures 3.4 to 3.6 respectively. The arithmetic average of the reverberation time T_{mf} in the 500 Hz, 1000 Hz and 2000 Hz octave bands are shown in Figure 3.7. The reverberation times presented in these figures are the averaged values in the three sets of measurements. Figures 3.4 to 3.6 illustrate that the reverberation times at different positions inside the classrooms are of similar values.

It is remarkable that the reverberation time can substantially be reduced if the classroom is finished with acoustic ceiling, see the reverberation times of the classrooms S5-A, S5-B and P4-A as shown in Figure 3.7, such that the reverberation times of these classrooms (except S5-A, it is marginally acceptable) comply with most of the international standards, i.e. RT < 0.6 s, except Italy where its requirement is generally too lenient. However, the requirement of RT < 0.4s as recommended by ASHA is too stringent. No classrooms in our study can comply with this recommendation.

It can be concluded from Figures 3.4 to 3.6 that the acoustic wall panel cannot provide an effective means to reduce the reverberant sound field inside the classroom as compared with the acoustic ceiling. The reverberation times of the classrooms treated with acoustic wall panel (i.e. P2-A, P2-B, P5-A and P5-B) all exceed the international limits, i.e. RT > 0.6 s. Excluding the extreme reverberation times recommended by ASHA, all of the classrooms without acoustical treatment have the reverberations times exceeding the international limit of 0.6 s. Five of them (i.e. S1-A, S2-A, S4-A, P3-A and P3-B) even have the reverberation times exceeding 1.1 s that are much higher than the international standards. Pictures 3.3 to 3.5 show the acoustic ceiling, and two different types of acoustic back wall panel of the measured classrooms.

The relative short in the reverberation times, where RT < 0.8 s, of the classrooms S1-B, S3-A, P1-A and P1-B compared to other untreated classrooms where RT > 1.1 s and even the classrooms installed with acoustic wall panel but RT > 0.8 s can attribute to the presence of bags, jackets, books and fabric stationery cases inside the classrooms, see Picture 3.1 and Picture 3.2 as a typical example. The presence of these sound absorptive materials can provide extra sound absorption inside the classroom that can substantially reduce the reverberation time of the classroom. This is possibly the reason why the classrooms have the same interior design but different reverberation times in our measurements. The occurrence of the fact that the reverberation times of classroom P4-A is shorter than that of classrooms P1-A and P1-B can reinforce the statement that the presence of bags, jackets, books and fabric stationery cases inside a classroom can reduce the reverberation time of the classroom even though the classroom is either acoustically treated or not. Though the human hearing system is not sensitive to the noise below 500 Hz and the frequency of almost all of our speech is distribute in 500 Hz to 2000 Hz, it is also noticeable to study the reverberation time in low frequency (< 500 Hz) as it may indicate the effect of mechanical noise to speech intelligibility.

3.4.4 Speech transmission index and rapid speech transmission index

The STI and RASTI at different positions inside the classrooms were obtained from averaging the three sets of measurements at position 1 to position 3 and presented in Figure 3.8 and Figure 3.9 respectively. According to the data acquired from the three positions inside the classrooms, position 2 have the highest value of STI and RASTI at most. Position 1 and Position 3 varies on the degree of STI and RASTI, though, Position 3 have a higher value in STI and RASTI than Position 1 comparatively. The STI and RASTI are dependent and sensitive to acoustical characteristics of the environment which include outdoor noise level, indoor ambient noise level, absorption coefficient of the environment surrounding to pricking balloon, especially in certain frequency bands.

Position 2 is placed at the cater-cornered centre where comparatively far from the outdoor noise source: windows and the doors. Effect of outdoor noise is the smallest among the three positions inside the classrooms. Apparently, the absorption powers of acoustic ceiling or carpet inside the classrooms are highly effective on middle to high frequency range, which are the most of sound energies of pricking balloon. Due to that absorption, the reverberation is decreased and therefore the highest value of STI and RASTI were found in position 2. Position 1 and position 3 are closed to the doors and the windows respectively. The STI and RASTI are lower in contrast as they are located in locations where easily affected by outdoor annoyance. In addition, position 3 is close to back wall that some of the classrooms with acoustic back wall panel can effectively decrease the effect of reverberation and diverse the pathway of speech transmission. Position 1 is located too close to the whiteboard and reflection of it induced masking and increase reverberation.

During the speech transmission inside the classrooms, the speech signal is changed due to the echo, reverberation and ambient noise that reduce the speech intelligibility. The measurement results show that speech intelligibility is dependent on the reverberation of the classrooms even though there have variations in ambient noise levels. Research showed that STI should be larger than 0.6 as a minimum requirement for effective communication and speech intelligibility [3, 5]. From the measured results, it is clear to observe that when the arithmetic average of the reverberation times (T_{mf}) are less than 0.6 s, see Figure 3.7, the STI and RASTI are larger than 0.7 for classrooms S5-A, S5-B and P4-A. However, when T_{mf} of the classrooms S1-A, S2-A, S4-A, P3-A and P3-B are larger than 1 s, the STI and RASTI are less than 0.6. The STI and RASTI of the remaining classrooms range from 0.6 to 0.7 while T_{mf} of the classrooms are between 0.6 s and 0.9 s. It is rational to have this observation as when the reverberation time is excessively long, the sound energy that does not reinforce the direct sound, in effect, becomes a masking sound for test signal, reducing speech intelligibility. For the above, reverberation time and STI are dependent and interact quantity to each other: larger RT induced to smaller STI due to masking effect and sound reverberant.

With reference to table 3.1, figure 3.8 and figure 3.9, the STI and RASTI in different classrooms vary even when the classrooms have same geometrical configuration. Therefore the acoustical configuration inside the classrooms plays an important role to those variations. Similar to discussion of reverberation time, the STI and RASTI is higher if the presence of bags, jackets, books and fabric stationary cases inside a classroom can induce a higher result in STI and RASTI, even though the classroom is either acoustically treated or not. The presence of bags, jackets, books and fabric stationery cases inside the classrooms increase the absorption that shorter the reverberation. Comparison of the STI and RASTI of the classrooms P1-A, P1-B, S-4A can strengthen the statement we discussed in section 3.4.3.

Besides, the area of absorption materials is an important factor for efficacy of improving the speech intelligibility. RT is an absorption dependent parameter and directly affects the STI and RASTI. The area of absorption materials is one of the important parameter in calculation of RT. It is therefore a larger area of absorption materials that can improve the speech intelligibility in the classrooms. From the results of carpeted classrooms (U1-A, U1-B, U1-C, U1-D), classrooms with acoustic ceiling (P4-A, S5-A, S5-B), classrooms with acoustic back wall panel (P2-A, P2-B, P5-A, P5-B), and the rest of the classrooms without acoustics treatment, the latter have the smallest STI and RASTI value since none of acoustics treatment inside the classrooms. Obviously the acoustic ceiling, carpet and acoustic back wall panel are commonly use for improve the speech transmission inside the rooms since the large area of absorption materials can be easily to attached to the wall or fixtures of the rooms. However, in comparison to the results for these three acoustic treatments, acoustic ceiling and carpet have better performance to strengthen the verbal communication. It is not difficult to interpret that the area of ceiling or area of floor is larger than area of back wall and leads the increase of area of absorption material, e.g. decrease the reverberation and increase the STI and RASTI. It is also understood the situation of occurrence of students' bags, jackets, books and fabric stationery cases inside the classrooms increase the absorption area that shorten the reverberation.

3.5 Conclusions

With references to national standards from thirteen countries and recommendation from three organizations, indoor ambient noise level and reverberation time are indispensable acoustical properties to assess the acoustical conditions for classrooms. Some of them specified noise criteria as one of the main parameter to assess the acoustical conditions for classrooms. The current acoustical performance of the classrooms in Hong Kong forms the initiative to compare the conditions with these international standards. During the study, field measurements have been carried out in the selected classrooms to acquire several acoustics quantities that related to speech intelligibility: indoor ambient noise level, outdoor noise level, noise criteria, reverberation time, speech transmission index and rapid speech transmission index. These quantities interact and depend on each other in deciding the speech transmission inside a classroom. Results showed the acoustical conditions in most of the classrooms studied are not satisfactory. Noisy environment and large ambient noise level, long reverberation time and large mechanical equipment induced noise also lower the STI and RASTI value. These unexpected results bring out the necessity to revise the standard design and to devise acceptable standards for classrooms in Hong Kong. This study also provides an indication for the interior designer on the acoustic environment of the classrooms after the installation of a proprietary material on the ceilings, floors or back walls in classrooms. The results of speech transmission are affected by absorption materials such as location of the materials, the area, and absorption coefficient of the materials, the absorption frequencies and other of their acoustical characteristics. The proprietary materials are designed to provide optimized sound absorption that will improve the acoustic environment of the classroom. Therefore, selection of materials inside a classroom is important for improving speech intelligibility that not only match the architectural and aesthetic requirements, but also their absorption and acoustical characteristics, construction and installation on the stage of design.

References

1. HOUTGAST, T. and STEENEKEN, H. J. M., “The modulation transfer function in room acoustics as a predictor of speech intelligibility,” Acustica **28**, pp. 66-73 1973.
2. HOUTGAST, T., STEENEKEN, H. J. M. and PLOMP, R., “Predicting speech intelligibility in rooms from the modulation transfer function I. General room acoustics,” Acustica **46**, pp. 60-72, 1980.
3. STEENEKEN, H. J. M. and HOUTGAST, T., “A physical method for measuring speech transmission quality,” J. Acoust. Soc. Am. **67**(1), pp. 318-326, 1980.
4. STEENEKEN, H. J. M. and HOUTGAST, T., “RASTI: a tool for evaluating auditoria,” Brüel & Kjær Technical Review 3, pp. 13-39, 1985.
5. HOUTGAST, T. and STEENEKEN, H. J. M., “A multi-language evaluation of the RASTI-method for estimating speech intelligibility in auditoria,” Acustica **54**, pp. 185-199, 1984.

6. VALLET, M. and KARABIBER, Z., “Some European Policies Regarding Acoustical Comfort in Educational Buildings,” Noise Control Eng. J. **50**(2), pp. 58-62, 2002.
7. VALLET, M., “Some European Standards on Noise in Educational Buildings,” International Symposium on Noise Control and Acoustics for Educational Buildings, Istanbul/Turkey, pp. 13-20, 2000.
8. Intelligibilité de la parole dans les salles de classe – Recommandation BIAP (Bureau International d’Audiophonologie) 09/10-4, Belgique – Février, 2003, http://www.biap.org/recom09-10-4.htm.
9. DIN 4109, Deutsches Institut für Normung: Sound insulation in buildings; requirements and testing, Berlin, 1989.
10. DIN 18041, Deutsches Institut für Normung: Acoustical quality of small to medium-sized rooms, Berlin, 2004.
11. Department for Education and Skill, Building Bulletin 93: Acoustic Design of Schools, 2003, http://www.teachernet.gov.uk/acoustics.

12. BFS 2002:19, Swedish Board of Housing, Building and Planning, Building Regulations BBR: Mandatory provisions and general recommendations, 2002, http://www.boverket.se/novo/filelib/arkiv10/bbr2002engelskavers27okt03.pdf.

13. ANSI S12.60-2002, American National Standard: Acoustical Performance Criteria, Design Requirements, and Guidelines for Schools, American National Standards Institute, New York, 2002.

14. CBD-92-F, Digeste de la construction au Canada: L'acoustique des salles publiques, Canada – publié à l'origine en mai, 1969, http://irc.nrc-cnrc.gc.ca/cbd/cbd092f.html.

15. GB/J 118, National Standard of Peoples Republic of China: Chinese Design Regulations for Civil Buildings, Chapter 4 School Building, National Technical Supervision Bureau, 1988.

16. AS/NZS 2107:2000, Standards Australia/New Zealand Standard: Acoustics – Recommended Design Sound Levels and Reverberation Times for Building Interiors, 2000.

17. World Health Organisation, Guidelines for Community Noise, 1999, http://www.who.int/docstore/peh/noise/Comnoise-4.pdf.

18. American Society of Heating, Refrigeration and Air Conditioning Engineers, Inc.: Heating, Ventilating and Air-conditioning Applications, 1999.

19. American Speech-Language and Hearing Association: Position Statement and Guidelines for Acoustic in Educational Settings, 1995, http://www.biap.org/recom09-10-4.htm.

Tables

School Code	Classroom Code	Design	Measured Volume / m^3	Mechanical Ventilation System	Acoustic Treatment
P1	P1 – A	standard	211	2 A/C & 2 FAP	-
	P1 – B				
P2	P2 – A	non-standard (no furniture)	235	4 A/C & 2 FAP	acoustic back wall panel
	P2 – B				
P3	P3 – A	standard	204	2 A/C & 2 FAP	-
	P3 – B				
P4	P4 – A	standard	210	4 A/C & 2 FAP	acoustic ceiling
P5	P5 – A	non-standard	224	4 A/C & 2 FAP	acoustic back wall panel
	P5 – B				
S1	S1 – A	standard	216	4 A/C & 2 FAP	-
	S1 – B				
S2	S2 – A	standard	211	2 A/C & 2 FAP	-
S3	S3 – A	standard	209	4 A/C & 2 FAP	-
S4	S4 – A	standard	215	4 A/C & 2 FAP	-
S5	S5 – A	standard	211	2 A/C & 2 FAP	acoustic ceiling
	S5 – B				
U1	U1 – A	non-standard	239	2 A/C & 2 FAP	carpet
	U1 – B				
	U1 – C		237		
	U1 – D				

Note:

[1] P indicates primary school, S indicates secondary school and U indicates university.

[2] A/C represents split type air-conditioner, FAP represents fresh air pre-conditioner.

Table 3.1 – Geometrical and acoustical details of the classrooms for

measurement.

Country & Organization	Indoor Ambient Noise Level / dB(A)	Reverberation Time / s	Noise Criteria
ASHA [19]	30	0.4 (500, 1k or 2k Hz)	20
Belgium [8]	40	0.4 ($200 \leq V \leq 1000$ m^3, averaged value for octave bands from 500 to 1k Hz)	-
France [6 – 7]	33, 38	$0.4 < T_{mf} < 0.8$ ($V \leq 250$ m^3) $0.6 < T_{mf} < 1.2$ ($V > 250$ m^3)	-
Germany [9 – 10]	35, 40	0.32 Log V - 0.17 (octave bands: 250 – 2k Hz)	-
Italy [6 – 7]	36	-	-
Portugal [6 – 7]	35	1 (octave bands: 125 - 250 Hz) 0.6 – 0.8 (octave bands: 500 – 4k Hz)	-
Sweden [12]	30	-	-
Turkey [6 – 7]	45	-	-
UK [11]	35	T_{mf}: 0.6 (Primary school) T_{mf}: 0.8 (Secondary school)	-
ASHRAE [18]	-	-	30
Australian / New Zealand [16]	35	0.4 - 0.5 (Primary school) 0.5 - 0.6 (Secondary school) (500 and 1 kHz)	-
WHO [17]	35	0.6	-
USA / ANSI [13]	35, 40	0.6 ($V \leq 283$ m^3), 0.7 ($283 < V \leq 566$ m^3) (500, 1k & 2 k Hz)	-
China [15]	Grade A: 40 Grade B: 50	0.9 ($V \leq 200$ m^3), 1.0 ($500 < V \leq 1000$ m^3) (500 Hz)	- -
Canada [14]	-	0.7 (30 students classroom)	-

Note:

[1] T_{mf}: Arithmetic average of the reverberation times in the 500 Hz, 1k Hz and 2k Hz octave bands.

[2] V: volume of classroom.

Table 3.2 – International standards, guidelines and recommendations

of different countries and organizations for the acoustical design of classroom.

Figures

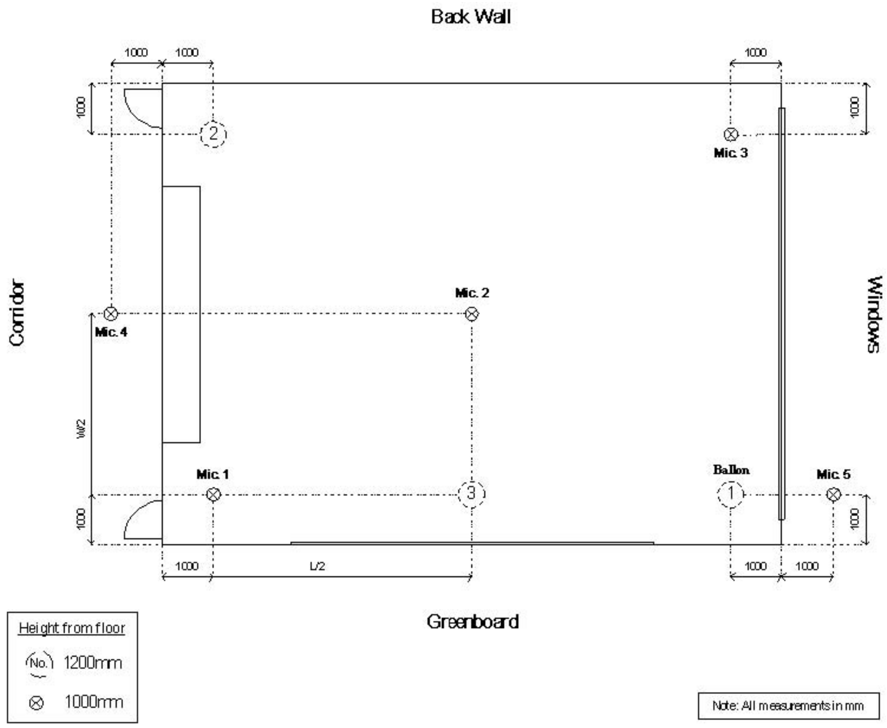
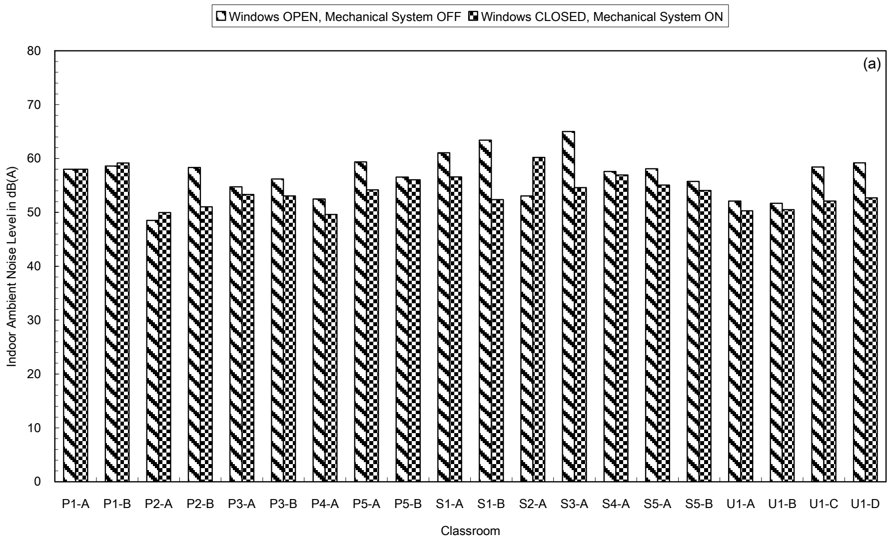


Figure 3.1: Details of experimental setup for measurement.



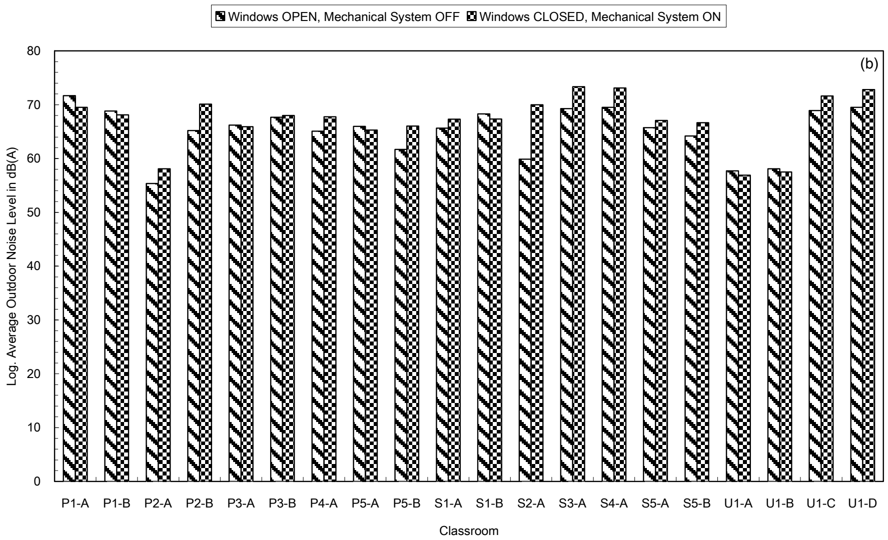


Figure 3.2 – (a) Indoor ambient noise levels of the classrooms under two different ventilation conditions in the measurements, (b) Logarithmic average of outdoor noise levels measured at position 4 and 5 where placed at corridor and 1m outside from the windows in the measurements.

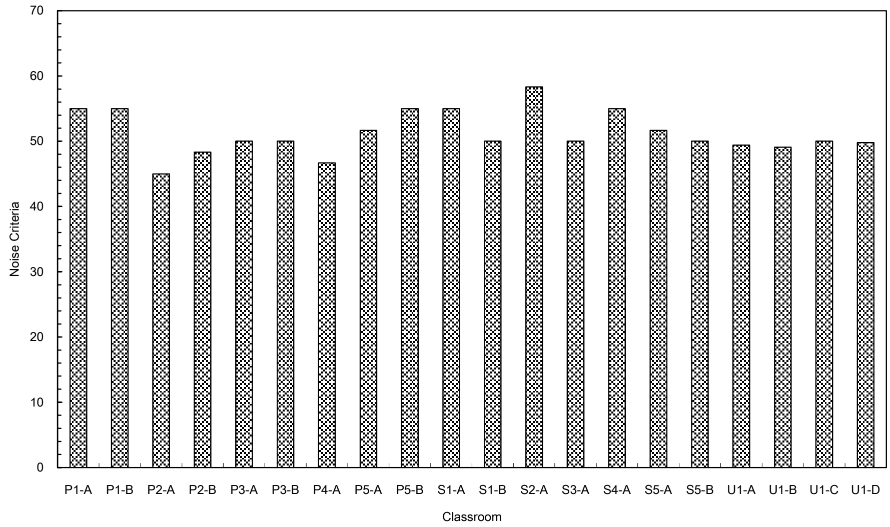


Figure 3.3 – Noise criteria of the classrooms under the condition of the windows and doors were closed but the mechanical ventilation system was switched on.

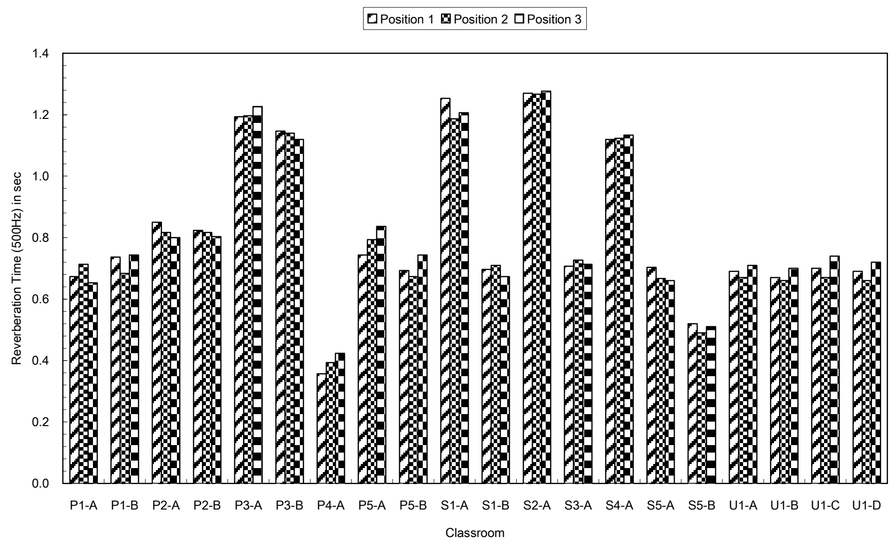


Figure 3.4 – Reverberation time at 500 Hz of position 1 to 3 inside the classrooms.

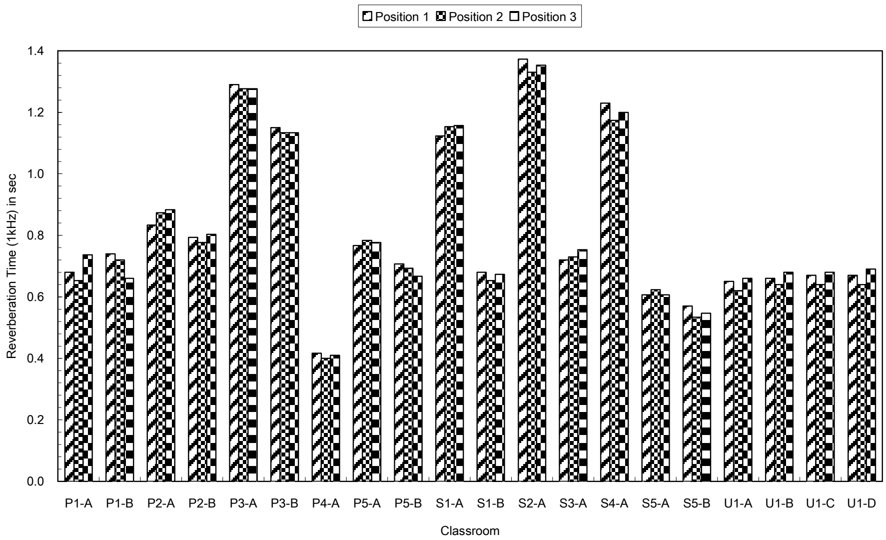


Figure 3.5 – Reverberation time at 1000 Hz of position 1 to 3 inside the classrooms.

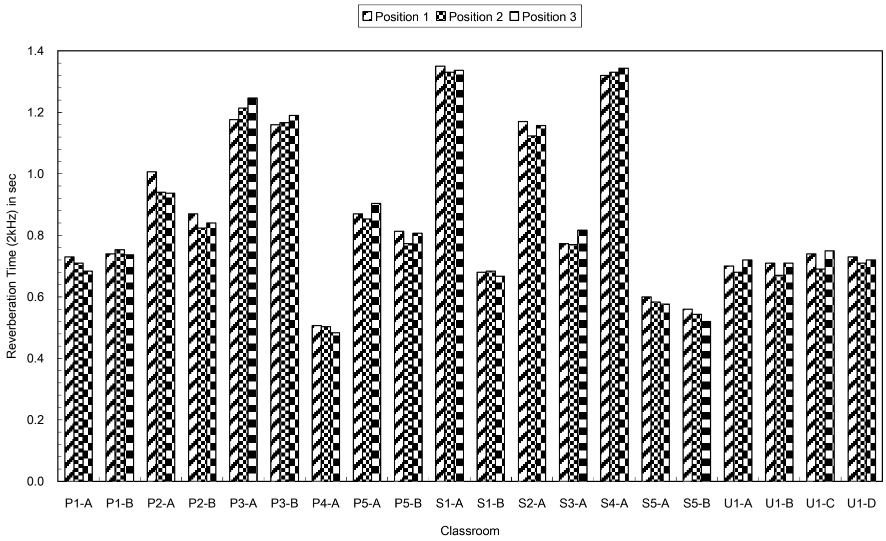


Figure 3.6 – Reverberation time at 2000 Hz of position 1 to 3 inside the classrooms.

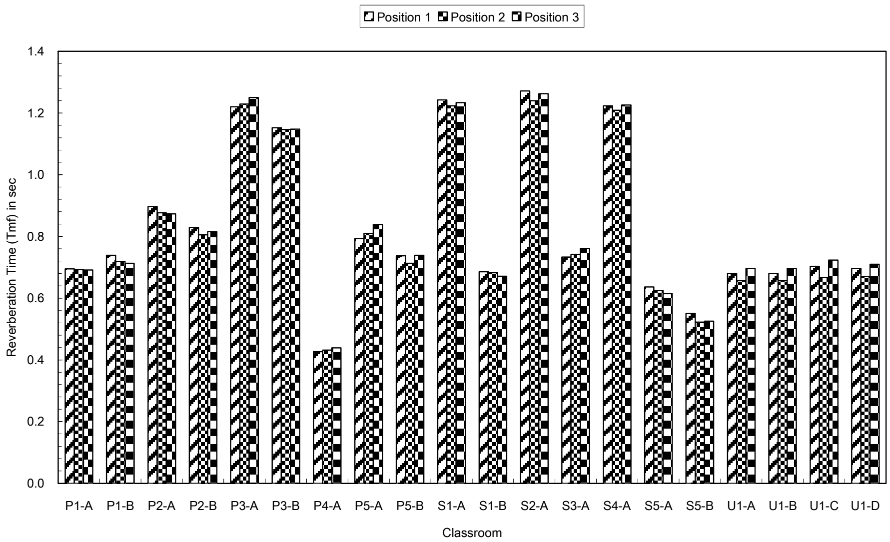


Figure 3.7 – Arithmetic average of the reverberation time (Tmf) in 500 Hz, 1000 Hz and 2000 Hz octave bands of position 1 to 3 inside the classrooms.

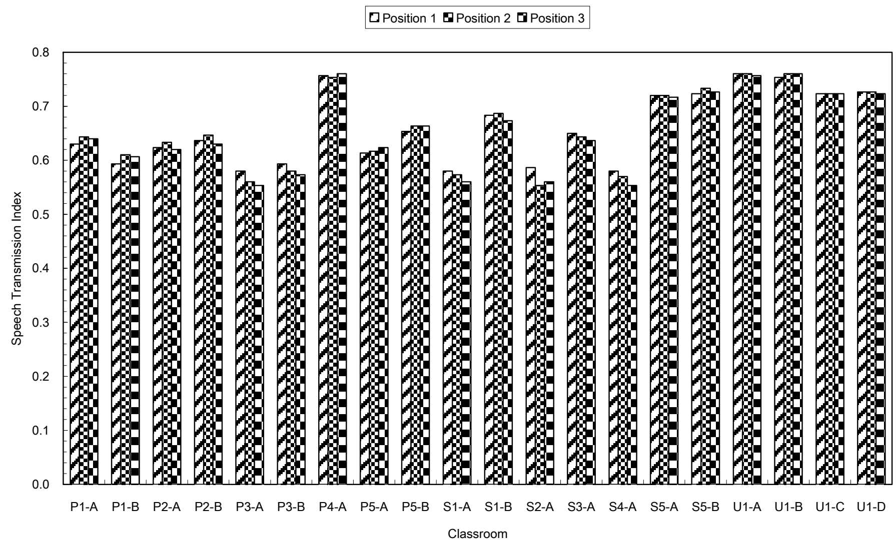


Figure 3.8 – Speech transmission index of position 1 to 3 inside the classrooms.

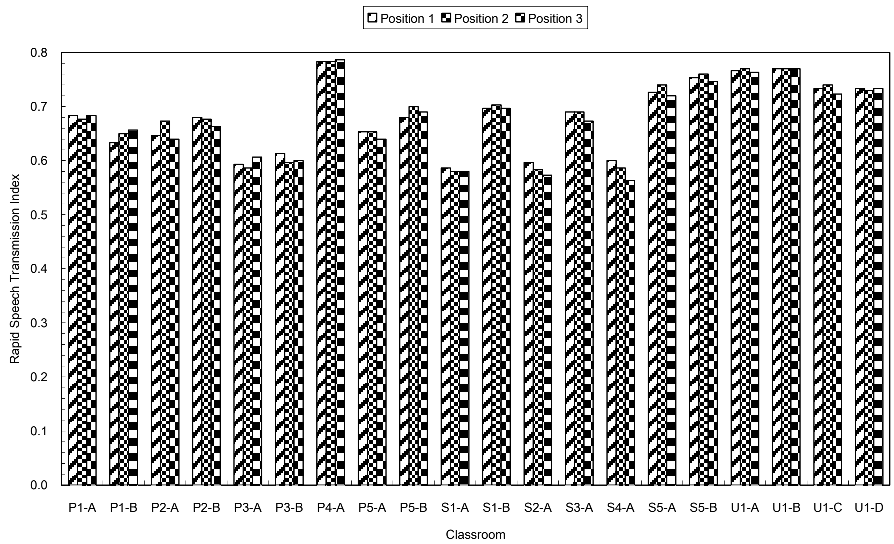


Figure 3.9 – Rapid speech transmission index of position 1 to 3 inside the classrooms.

Pictures

Picture 3.1 – Details of classrooms P1-A and P1-B, which with presence of bags, jackets, books and other furniture.

Picture 3.2 – Details of classrooms S4-A and S4-B, which without presence of bags, jackets, books and other furniture.

Picture 3.3 – Acoustic ceiling detail of classroom S4-A.

Picture 3.4 – Acoustic wall panel of classroom P2-A.

Picture 3.5 – Acoustic wall panel of classroom P5-A.

Chapter 4
Acoustics simulations and predictions for improving the speech intelligibility in classrooms

4.1 Essential needs of computerize simulation software

Teachers and students have meaningful communication in classrooms. Therefore, ensure the speech intelligibility is essential to maintain the effectiveness of verbal communication. Investigation of an effective acoustics communication system in the classroom can provide an excellent sound environment. This will be an advantage to both teachers and students psychological and physical health, and strengthen the interaction of teaching and learning.

The superior request of speech intelligibility in the classroom should have short reverberation time, enough energy from signal source, even distribution of sound field, no acoustic interference, minimal diffusion or echo effect to signal source, and no environment interference or noise, etc. To reach these requirements, it is necessary to consider the acoustic condition and noise control during design stage of the classroom [1]:

1. The acoustics design of classroom has to cohere to the architectural and interior design plans. With selection on structural configuration and proper reverberation time, the decision on arrangement of reflection and absorbing materials is important. This will enable the sound to be delivered perfectly to the students. Sometimes the address system may need to be set up.
2. To control the noise and make the acoustics more successful, thorough discussion and consideration has to be made on urban planning for construction design. Material selection is not only important in architecture, but also in the acoustics. For example the absorptive materials, noise abatement precaution, noise reduction and insulation from environment all need to be taken into consideration.

However, noise control engineering or acoustical design for the classroom is not an easy topic. We can calculate the acoustics parameter during the acoustics design stage, but the border condition is complicated and its calculation is very time consuming. The results cannot be guaranteed to be accurate, so it is not a widely practiced application.

Typically, the architect will choose and install the absorptive materials and their size by experience in architectural acoustics design. The sound field and acoustic effect can not be defined in the classroom, which needs to be tested after construction. Renovation is the only way to modify the classrooms with poor acoustics conditions. However, the amount of modifications are very limited and hard to reach the best results for previously built classrooms since most of the acoustics treatment has been paved and the related component has been set up. Any change and modification will cause time consuming construction, which can be a huge investment loss. The acoustic conditions inside a classroom are sensitive to absorptive surrounding material with different specification, size, and installation. Field research is too late to being in order to tell what level can be reached for speech communication in a specific classroom. During the preliminary design stage acoustic estimation is commonly used as a synthetic consideration for the parameters to achieve an expected range of speech intelligibility. Many acoustics launched types of software can settle the problem effectively [2].

Using the acoustics simulation software on architecture has been studied internationally and well developed commercial software has been launched [3]. These softwares can simulate the calculation of acoustics parameters accurately and imitate a near real result which can enhance the consideration for design and build.

RAYNOISE, is one of the acoustics software which will be introduced and instantiate its application on this study for classroom acoustics.

4.2 Introduction of RAYNOISE

RAYNOISE is developed by Belgium LMS who is one of the leading developers of computer-aided engineering software for computational vibro-acoustics and engineering design analysis [4]. It is a computer-aided acoustic modelling and analysis system for the application in different aspect. The architecture and theory is based on ray-tracing method to predict the sound field produced by multiple sources at different locations either in closed, open or partially open spaces. Acoustical parameters and acoustics performance inside the space under a complex situation and environment such as: different acoustics features, can be simulated and modelled accurately. During the simulation consideration of diffusion, diffraction, reflection, etc. are necessary to obtain the maximum results.

This software not only has an excellent interface but also has high precision and speed in calculation for speech intelligibility [3]. Their features are noticeable and user friendly. Unlimited amount of calculation of the model’s surface can maximize the simulation for forecasting the sound field. Similarly, the amount of absorptive materials that can be used depends on how many surfaces there are for the model. There are unlimited conditions for calculation of model. Since it is based on ray-tracing method, the user can customize the reflect order and its ray number. With this characteristic, unlimited sound source can be chosen for the model and the selected source for results. The theory on simulation calculation for absorption is based on the Sabine equation. The reverberation can then easily find the variation of absorption coefficient of the proprietary materials and surface area of its installation. Its open database can easily be modified discretionarily and receive the text data file.

4.3 Simulation of selected classroom

4.3.1 Objectives

For improving the acoustic environment of the existing primary and secondary classrooms in Hong Kong, a specific classroom has been chosen for a mock-up design. The acoustic environment in the classroom was assessed by a comprehensive numerical simulation of the reverberation time, and the voice levels at different positions for a given interior design of the classroom.

The simulations were conducted using a computer-aided software package, RAYNOISE Rev. 3.1.

The results of simulations gave an insight for the acoustic environment of the classroom. The simulations were conducted after the installation of an overhead reflector, and proprietary sound absorption materials on the ceiling and the back wall in the classroom. The overhead reflector and proprietary materials aim to provide improved signal-to-noise ratios and optimised sound absorption respectively. The goal is to increase the speech intelligibility in the classroom.

4.3.2 Details of the selected classroom

The geometry of the classroom is based on the architectural drawings and a 3D drawing of the corresponding classroom measured. In all the numerical modelling, the sound source is assumed to be located at 1.5 m above the floor and 1 m in front of the middle of the white board. The source location and details of the classroom are given in the isometric view of the classroom as shown in Figure 4.1. The sound source is simulated by an omni-directional speaker with the source strength equivalent to a human voice as built-in the database of RAYNOISE. In the present study, two thousand sound rays were generated and traced with a maximum being 40 orders of reflections.

The acoustical properties of surfaces, such as the door, floor, windows and wall panels etc., inside the classroom are obtained either from a well-known reference book of engineering noise control [6] or the built-in database provided by RAYNOISE since the general characteristics of common materials used and quoted from specification of current commercial products. The acoustical properties of the absorptive ceiling and absorptive wall panel are based on the commercially available brand ‘Ecophon Focus^{TM} E’ and ‘Ecophon Wall Panel^{TM} C’ respectively. The locations of the absorptive ceiling ‘Ecophon Focus^{TM} E’, absorption back wall panel ‘Ecophon Wall Panel^{TM} C’ and overhead reflector are shown in Figure 4.2. Figure 4.3 give a plan view and a front view of the classroom in which the orientation of the overhead reflector is shown to be inclined at 45^{o} to the vertical axis (measured along the height of the wall). The overhead reflector is modelled with a width of 1.13 m and a length of 6 m. The sound absorption coefficients of the absorptive ceiling, absorptive wall panel and other surfaces in the classroom at the octave frequency bands are given in Table 4.1 for ease of reference. The acoustical properties of other surfaces inside the classroom are obtained either from Sharland [6] or the built-in material database provided by RAYNOISE.

In this study, the reverberation time and voice levels at different positions are used as acoustic parameters for assessing the acoustic performance of the classroom. The receiving points are taken as 1.2 m above the floor in all the calculations. The reverberation time is especially useful as it relates to the noise reduction given by an absorbing treatment in an enclosed space. It is because when a sound source operates in an enclosed space, the level to which reverberant sound builds up, and the subsequent decay of reverberant sound when the source stopped, is governed by the sound-absorbing characteristics of the boundary surfaces and objects within the space. The voice level is a crucial parameter for the calculation of signal-to-noise ratio which is directly related to the speech intelligibility at a particular location in the classroom. Air absorption inside the classroom was taken into account in the acoustic modelling and simulation. Besides, acoustical characteristics of reflection materials will affect the sound field, therefore there is an analysis of reflection material selection which according to simulate different acoustics specification of materials.

4.3.3 Simulation results and analysis

In a classroom with absorptive ceiling, different angle of reflection panel will affect the reverberation time and sound intensity result. Figure 4.4 and Figure 4.5 show the respective simulated voice levels and reverberation times at 1 kHz inside the mock-up classroom with the installation of absorptive ceiling and overhead reflector at different oblique angles to the vertical axis. The simulation results reveal that the voice levels at the seating zone can be increased by about 1 dB (ranging from 47.18 dB to 51.77 dB compared to the range between 45.90 and 50.95 without the reflector) with the presence of overhead reflector at an oblique angle larger than 45^{o}. On the other hand, the reverberation time can be minimized (ranging from 0.82 s to 0.87 s without the reflector compared to the range between 0.72 s to 0.76 s) with the overhead reflector inclined at 45^{o} to the vertical axis.

From the results showed in Figure 4.4 & Figure 4.5, regarding to the signal level, it keeps slightly increasing nearby the front of sound source when added the reflector at the overhead and change the angle from 15° to 75°. However, the sound level nearby the back wall is increasing by increasing the inclined angle when oblique angle 60° is adjusted, and is reducing when the oblique angle is 70°. After comparison of the results, the signal level distribution is the best and even when the reflector set as 60°. Besides, the reverberation time will reduce when the angle increased of the reflector. When 45° in angle, the distribution and the maximum value of reverberation time is relative small but it has been increase when the angle keep increasing. The simulation results for the octave band frequencies of 500 Hz and 2 kHz can reach the same conclusions but not shown here for brevity.

In the classroom with absorptive ceiling and acoustic panel at the back wall, different angle of reflection panel will affect the reverberation time and sound intensity result. The simulation results of voice levels and reverberation times inside the mock-up classroom with the installation of proprietary materials on the ceiling and back wall, and overhead reflector at different oblique angles to the vertical axis are shown in Figures 4.6 and 4.7 respectively. Similar conclusions can be drawn from the simulation results that the voice levels at the seating zone can be amplified from the range between

43.92 dB and 50.42 dB (without an overhead reflector) to the range between 44.77 dB and 51.26 dB when an overhead reflector is built at an oblique angle larger than 45^o. The reverberation time can be reduced from the range between 0.6s and 0.67s (without an overhead reflector) to the range between 0.59s and 0.64s with the overhead reflector at an oblique angle 45^o to the vertical axis. However, the optimum angle of reflector for the optimal reduction on reverberation time in this case is 75^o instead of 45^o such that the reverberation time can be reduced to the range from 0.56 s to 0.6 s.

Regarding to the signal level on Figure 4.6 and Figure 4.7, when added the reflector at the overhead and change the angle from 15° to 75°, the signal level is amplifying everywhere inside a classroom despite it is nearby the sound source or far away from the sound source. When the oblique angle set at 75°, the signal is maximum in level. According to the distribution of sound level inside a classroom, the optimum angle of reflector is 60° which shown in Figure 4.6.

Regarding to the reverberation time, it was changed volatile by the angle of reflector. Under different condition of the reflector's angle, the reverberation time was increased when the reflector with oblique angle 15° and 30°. Nevertheless, the reverberation time was obviously reduced when the oblique angle of reflector adjusted to 45°, increased when adjusted to 60° but decreased when further adjusted to 70°. Compare with all the result, the shortest reservation time is in oblique angle 60° of reflector. As shown in Figure 4.7, mostly the range around 0.58 second in the classroom but the highest value of 0.6 second was only in small size. Comparatively, the intelligibility of this classroom is the highest.

It is interesting to note that the reverberation time can be largely reduced with the installation of proprietary materials on the back wall in addition to on the ceiling as evidenced in Figure 4.5 and Figure 4.7. The reverberation times as shown in Figures 4.5 (a) and 4.7 (a) are found to drop significantly from the range between 0.82s and 0.87s to the range between 0.6s and 0.67s. However, the sound level can be slightly impaired with the installation of proprietary materials on the back wall in addition to on the ceiling as evidenced in Figure 4.4 and Figure 4.6. The sound level as shown in Figure 4.4 (a) and Figure 4.6 (a) are found to drop significantly from the range between 50.95 dB and 45.90 dB to the range between 50.42 dB and 43.92 dB. Similar conclusions can be drawn for the octave band frequencies of 500 Hz and 2k Hz.

Selecting the reflector materials will affect the sound quality in classroom [7]. The reverberation time will be reduced when the absorptive parameterises of reflector is larger. Though, the average sound pressure level is opposite which mean the reflective result will be better if the absorption coefficient is smaller, the average is larger then the sound pressure will be lesser and sound distribution will be average. Due to the smaller size of reflector, it will not be affect the reverberation time and general sound pressure but can see the influence of speech intelligibility. Alcons is one of the evaluation indexes of speech intelligibility. It is percentage articulation loss of consonants and is derived from direct-to-reverberation and early decay time. It expresses loss of consonant definition that consonants have a significant effect than vowels in speech intelligibility. Intelligibility will be increased if Alcons is smaller. Generally, the prefer level of Alcons is 0% to 7%; 7% to 11% is good; 11% to 15% is average; 15% to 18% is below average and above 18% is failure or unacceptable. With references to Table 4.2, selected aluminium micro-perforated panel with added glass wool board, which is the largest in acoustics reflective coefficient, the Alcons is the least in value. The percentage in 32% is good and average percentage is 68%. With paved reflector which has the smaller reflection effect, the Alcons will be smaller. In general, using the aluminium micro-perforated panel as the reflector or reflective materials will have a better result.

According to the above analysis, the reflector does not only have a good reflectivity but also affects the acoustics parameter such as reverberation time and average sound pressure level. In order to achieve the good acoustics condition, it may need to select the panel (wood or metal) with damping materials and also need to avoid the vibration when selected metal panel is used.

Micro perforated absorber (MPA) was developed by Maa in 60's [8]. He developed the absorber when he was prepared to apply it under adverse conditions. Absorber means fixing the micro perforated plate on the solid material and the hole is deci-millimetre. It was high in acoustic impedance and low acoustic resistance which became a wideband absorber. MPA is a simple structure and its absorptive characteristics can be calculated precisely. Unlimited the panel materials, it can be paper board, plastic, metal board or even film etc. MPA is developing rapidly and apply on many different categories.

Aluminium micro perforated panel have an acoustics characteristic as a wideband absorber. The vibration absorbs the peak value while moving to the low frequency when the cavity is expanding but the peak value of absorption coefficient was reduced. According to the cavity expanded and peak value increased which have enhanced the wideband absorption coefficient on aluminium micro perforated panel.

On the previous section, acoustics simulation and analysis have been performed for different materials. It is proved that the advantage of aluminium micro perforated panel (sometime with the glass wool board) is incomparable to those of other materials and even as a reflector or overhead absorber. Therefore, we are going to discuss the application of aluminium micro perforated panel.

Four acoustics measures are suggested in a classroom: 1) overhead mineral wool and aluminium micro perforated reflective panel; 2) overhead mineral wool and aluminium micro perforated reflective panel with additional 50 mm glass wool reflective board; 3) overhead aluminium micro perforated panel and aluminium micro perforated reflective panel; and 4) aluminium micro perforated panel and overhead 50 mm glass wool board with additional 50 mm glass wool reflective board. The surface of aluminium micro perforated panel is 1 mm in thickness with an aperture is d = 0.8 mm and perforated percentage is p = 1%.

This type of micro perforated panel is good in sound absorption of mid or low frequency and the absorption coefficient is small in high frequency which can satisfy the absorptive requirement in classroom and match the frequency characteristic of reverberation time. If the perforated percentage p is 1%, then 99% of smooth surface can be used on sound reflection and lead the classroom to have enough of early-arriving energy which enhanced the speech intelligibility. Therefore, the overhead aluminium micro perforated panel and aluminium micro perforated reflective panel will cause a trivial result on improvement. With addition 50 mm glass wool board in the micro perforated panel, it enhanced the sound absorption so the result will be comparatively better among those suggestions. Unfortunately, these two options are expensive. With consideration to the cost, overhead mineral wool and aluminium micro perforated reflective panel with addition 50 mm glass wool reflective board is the best mean. Due to the absorption coefficient of mineral wool in the frequency range being average, 0.42 to 1.00, it can satisfy the noise reduction requirement in classroom that the average of absorption coefficient is 0.71, sound reduction parameter is 0.75. With aluminium micro perforated reflective panel and added 50 mm glass wool board, the classroom can gain from early-arriving energy when increasing the sound absorption. It can reflect the voice from teacher to front, middle and back seating zone and enhanced the sound pressure of classroom, also average the distribution of sound field. Furthermore, the cost is comparatively cheaper among those suggestions that it is the best option for acoustics treatment by use of MPA.

By controlling the reverberation time between 0.6 s to 0.7 s in a classroom, higher speech intelligibility is reached [9]. According to the mock up result, the effect of absorptive ceiling is not perfect and the reverberation time is around 0.8 s but with the absorptive panel at the back wall, the reverberation time reduced to 0.6 s to 0.7 s. This can obviously improve the speech intelligibility in the classroom. Beside, if the reflective panel above the sound source is in the oblique angle to 75°, it induces a better result and enhanced the speech intelligibility, see Figure 4.6 (f) and Figure 4.7 (f). Usage of reflective panel should be considered as guideline in design of the classroom. The reflective panel can be hanged under the ceiling of a classroom or paved at two opposite side of blackboard. With the exception of satisfying the construction requirement, the reflective panel in the classroom should be installed at the lowest location. Therefore, it can reduce delivery time of direct voice and reflective time to the listeners. The oblique angle and the location of reflective panel depend on the receiver requirement at the back of the classroom. The reflective panel should be wide enough to enhance the back seating zone in such a way that the listener can hear the speech clearly by the reflective sound in very short response. In order to avoid the reflective sound interfere by diffraction of border of the panel, the length of reflective panel cannot be less than 3 m. The absorption coefficient of the panel should be small in all the frequency which can enhance sound absorption in different frequency and also provide enough energy of early-arriving to enhance the speech intelligibility in classroom.

4.3.4 Comparison of simulation and field measurement

In previous chapter, selected classrooms for field measurements on acoustical conditions about indoor ambient noise, outdoor noise, noise criteria from mechanical ventilation systems inside the classrooms, reverberation time, speech transmission index and rapid speech transmission index have been studied. In comparison of the results measured, it is found that various acoustical features inside the classroom are related to speech intelligibility. The speech transmission is affected by the nature of the materials, installation size and location, absorption coefficient and absorption frequencies. It is the vital consideration for those factors during the design stage of the classrooms. In this chapter, the acoustics simulation and modelling software for prediction on selection of proper acoustics measures inside the classrooms were investigated and studied. With the presence of proper acoustical treatments inside the classrooms, the consequence effect can be minimized due to reverberation, and therefore the speech intelligibility improved.

Refer to Chapter 3, it can be concluded from Figure 3.4 to Figure 3.6 that the acoustic wall panel cannot provide an effective means to reduce the reverberant sound field inside the classroom as compared with the acoustic ceiling. The argument is remarkable if only one acoustics treatment is applied to the classroom, acoustics ceiling have better performance in reducing the reverberation time than acoustics back wall panel. However, quality of verbal communication cannot be improved only by shortening the reverberation time.

According to the simulation results of a classroom with same geometrical configuration, it is evident that the reverberation can be significantly decreased with the installation of proprietary materials both at the ceiling and on the back wall. Figure 4.5 (a) and Figure 4.7 (a) illustrate that the reverberation times have a significant drop from the range between 0.82 s and 0.87s to the range between 0.6 s and 0.67 s. It is suggested to install the acoustics ceiling during the early construction stage as the reverberation can be improved thereafter by easily installation of the acoustics back wall panel if the classroom is unsatisfactory in acoustics conditions. Therefore, an optimization on combination of acoustics materials to improve the acoustics quality of classrooms should carry out in future study.

4.4 Conclusions

By large, acoustic treatment is not commonly used, and it seldom takes into consideration the design of a classroom. Although some of the architects used for improvement of acoustical condition of the classroom designed, the selection and size are based on experience but not by fact or data. From the results of field measurements in previous chapter and the simulation in this chapter, the need of proper acoustics treatment has been evidently identified. With the presence of acoustics ceiling, the reverberation time is reasonably shorten, however, the reverberation is significantly reduced with presence of both acoustics ceiling and acoustics back wall panel. Besides, other effective acoustics treatments like reflective panels are discussed on their performance. Several options and combination on the usage of aluminium micro perforated panel are recommended. Though, the acoustics treatment with fibre materials should be avoided in classrooms with respect to health consideration.

The popularity and the applicable features of acoustical simulation and modelling software are further studied by RAYNOISE. Due to its thorough theoretical background and architecture behind the program, complex acoustical configuration can be simulated and predicted the impact of variation of parameters. Various situations on the acoustics measures are applied to the program for simulation. It is also cost effective, practical and user friendly for consultants and educators to use as preliminary investigation on different cases during the planning stage. However, scattering effect and boundary consideration in classrooms may vary and complex, it may limit the accuracy of results from RAYNOISE. Conclusive results have been observed, indicating that presence of absorptive material inside a classroom is essential for improving the speech intelligibility since its absorptive characteristics in shorten the reverberation, mainly the reverberant frequencies.

References

1. PENG, J. X. *et al.*, “Prediction of classroom acoustic performance based computer simulation”, Audio Engineering 12, 2004.

2. PENG, J. X. *et al.*, “The application of acoustic simulation software in speech intelligibility evaluation”, Audio Engineering 11, 2002.

3. VORLÄNDER, M., “International round robin on room acoustics computer simulations”, 15th International Congress on Acoustics, Trondheim, Norway, pp. 577-580, 1995.

4. Manual of RAYNOISE version 3.1, LMS, Belgium.

5. ZHANG, X. H., “The application of acoustic simulation software in the noise prediction and evaluation”, China Academic Journal of Sound and Vibration 1, 2002.

6. SHARLAND, I., “Woods practical guide to noise control”, Colchester, England, Chap. Woods Acoustics, 1972.

7. JI, X. R., “Research of acoustical environment in classroom”, Taiyuan University of Technology, Thesis of Master of Philosophy, 2006.

8. MAA, D. Y., “Potential of micro perforated panel absorber”, J. Acoust. Soc. Am. 104, pp. 2861-2866, 1988.

9. ANSI S12.60-2002, American National Standard: Acoustical Performance Criteria, Design Requirements, and Guidelines for Schools, American National Standards Institute, New York, 2002.

Tables

Material	Frequency Band, Hz					
	125	250	500	1k	2k	4k
Absorptive Ceiling (Ecophon Focus™ E)	0.45	0.8	0.9	0.85	0.9	0.9
Absorptive Wall Panel (Ecophon Wall Panel™ C)	0.2	0.7	1	1	1	0.95
Floor	0.05	0.05	0.1	0.1	0.05	0.05
Wooden Panel	0.1	0.07	0.05	0.05	0.04	0.04
Doors	0.3	0.25	0.2	0.17	0.15	0.1
Windows	0.06	0.04	0.02	0.02	0.02	0.02
Plastering Wall	0.03	0.05	0.04	0.06	0.08	0.08
White Board	0.14	0.11	0.1	0.06	0.05	0.05
Tackboard	0.05	0.03	0.35	0.4	0.5	0.5
Overhead Reflector	0.14	0.11	0.1	0.06	0.05	0.05

Table 4.1 – Sound absorption coefficients of the absorptive ceiling, wall panel and other surfaces.

Material	RT, sec		SPL, dB		Alcons, %	
	500Hz	**1000Hz**	**Average**	**Deviation**	**Good**	**Average**
Baffle plate	1.04	0.95	54.7	1.60	14	86
Batten board	1.07	0.99	54.9	1.50	14	86
Fibreglass	1.07	0.99	54.9	1.60	14	86
Aluminium micro perforated panel	0.98	0.92	54.5	1.60	18	82
Aluminium micro perforated panel w/ glass wool board	0.91	0.85	54.1	1.70	32	68

Table 4.2 – Acoustical characteristics of baffle board with different material.

Figures

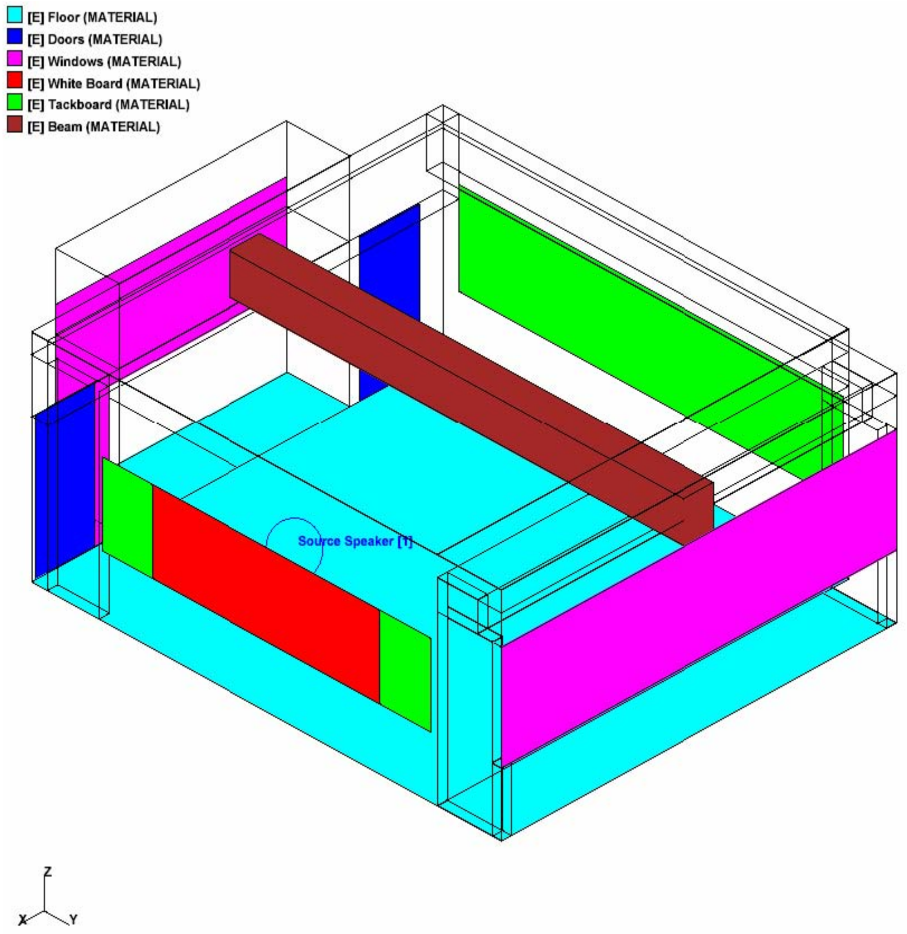


Figure 4.1 – Isometric view showing the details of the mock-up classroom.

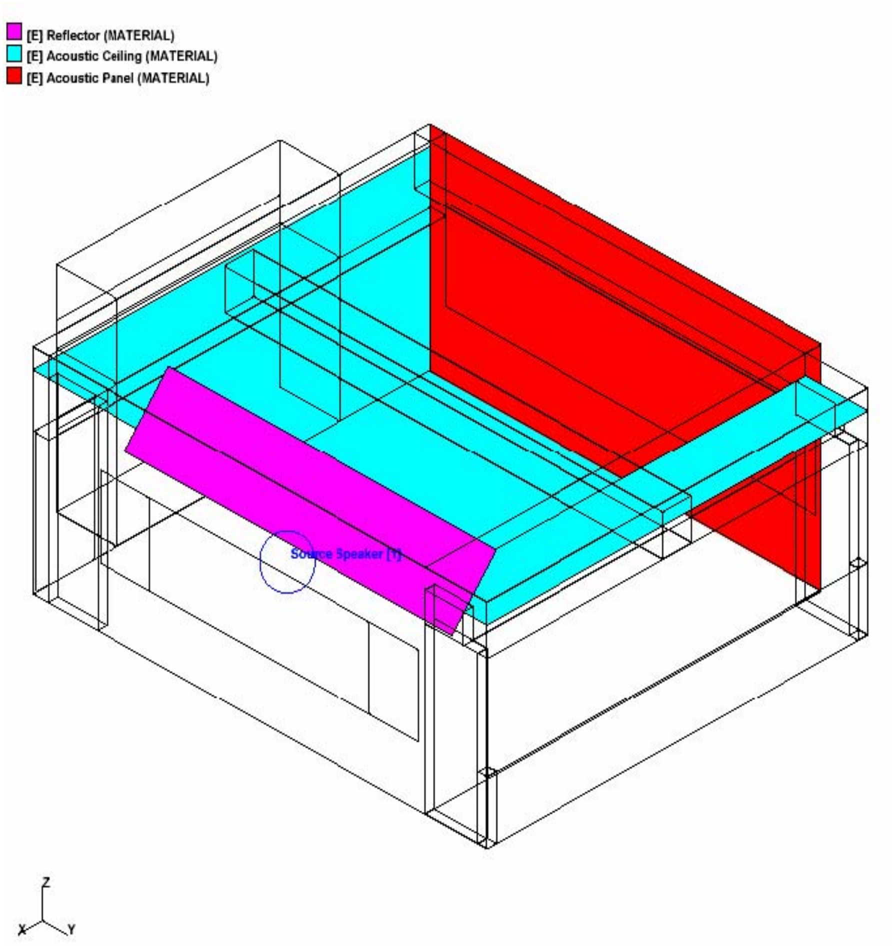
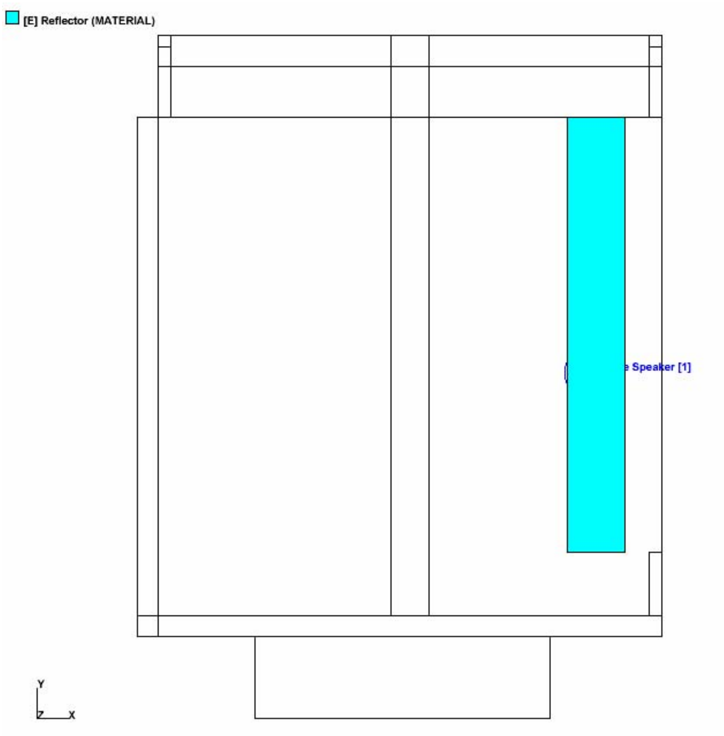
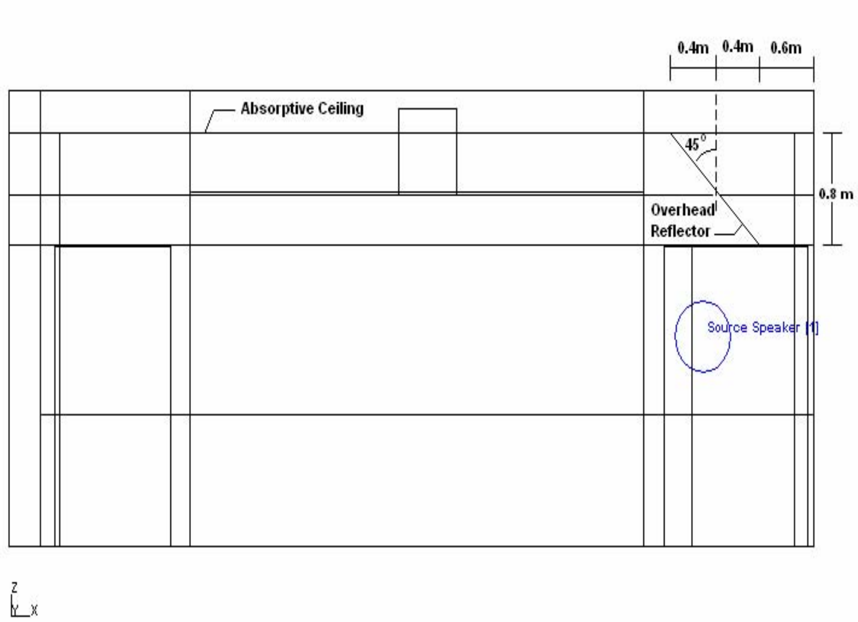


Figure 4.2 – Isometric view showing the locations of the overhead reflector, absorptive ceiling and back wall panel.



Plan View



Front View

Figure 4.3 – Diagram showing the orientation of the overhead reflector at an oblique angle 45° to the vertical axis.

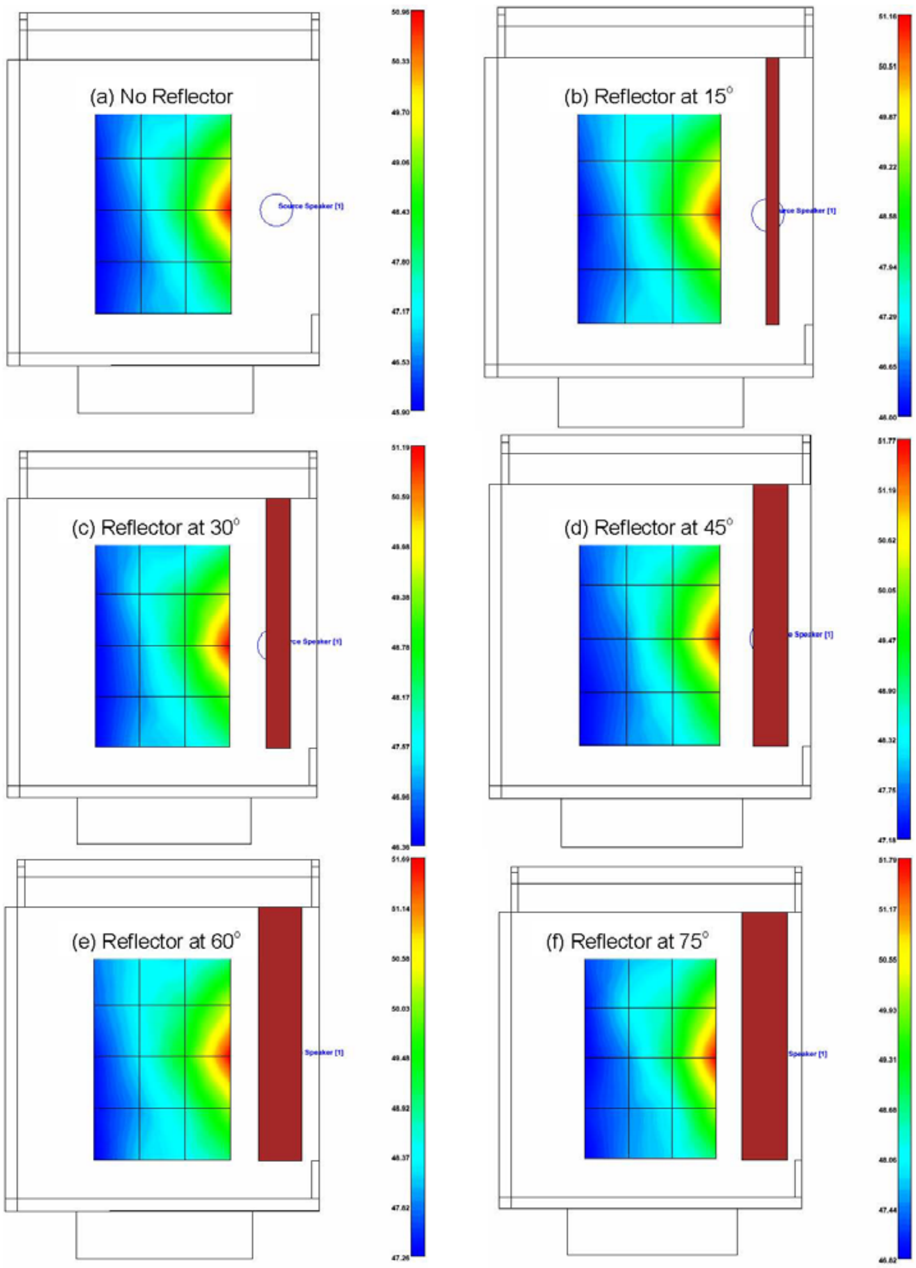


Figure 4.4 – Simulated voice levels in dB at 1k Hz inside the mock-up classroom (with absorptive ceiling) for reflector at different oblique angles.

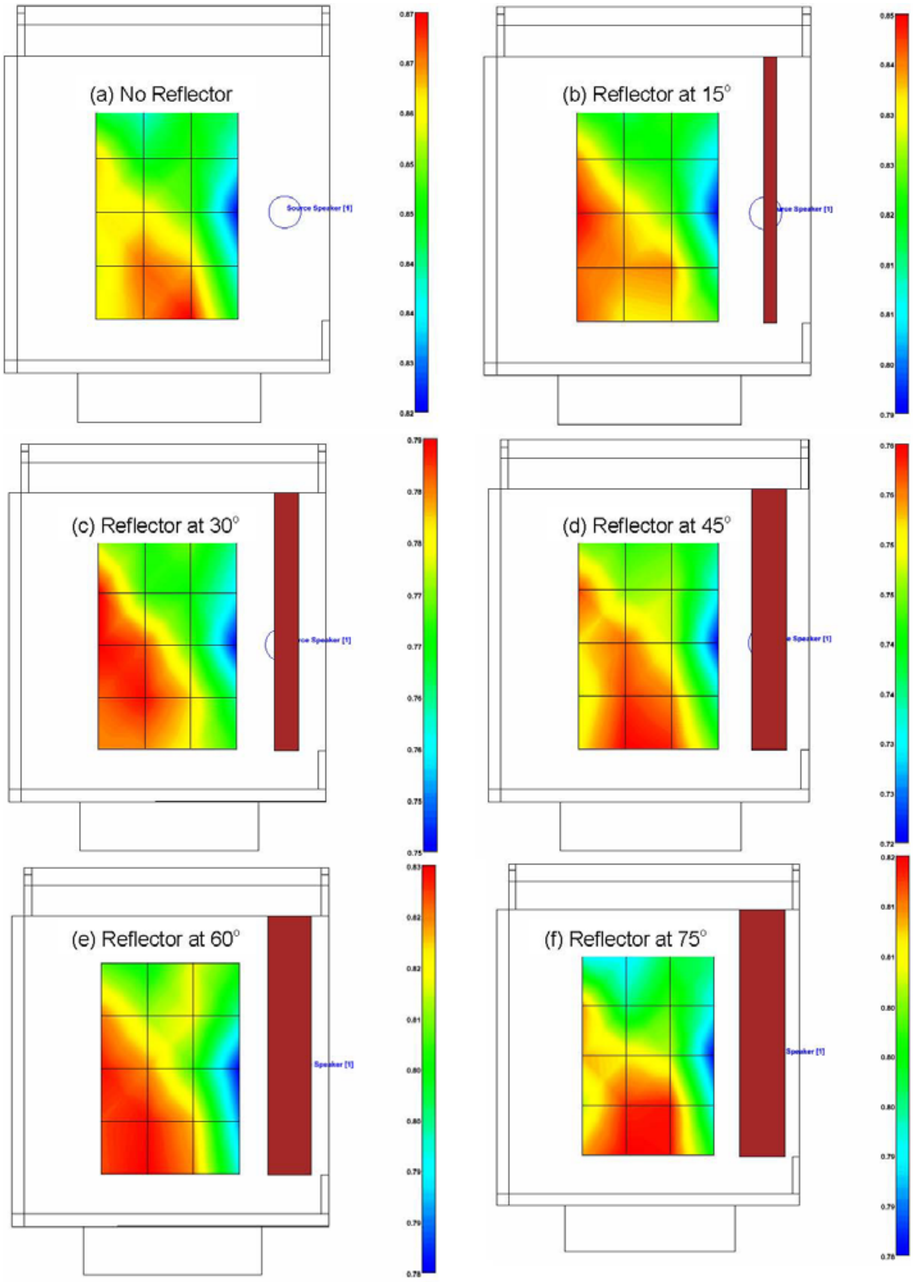


Figure 4.5 – Simulated reverberation times in second at 1k Hz inside the mock-up classroom (with absorptive ceiling) for reflector at different oblique angles.

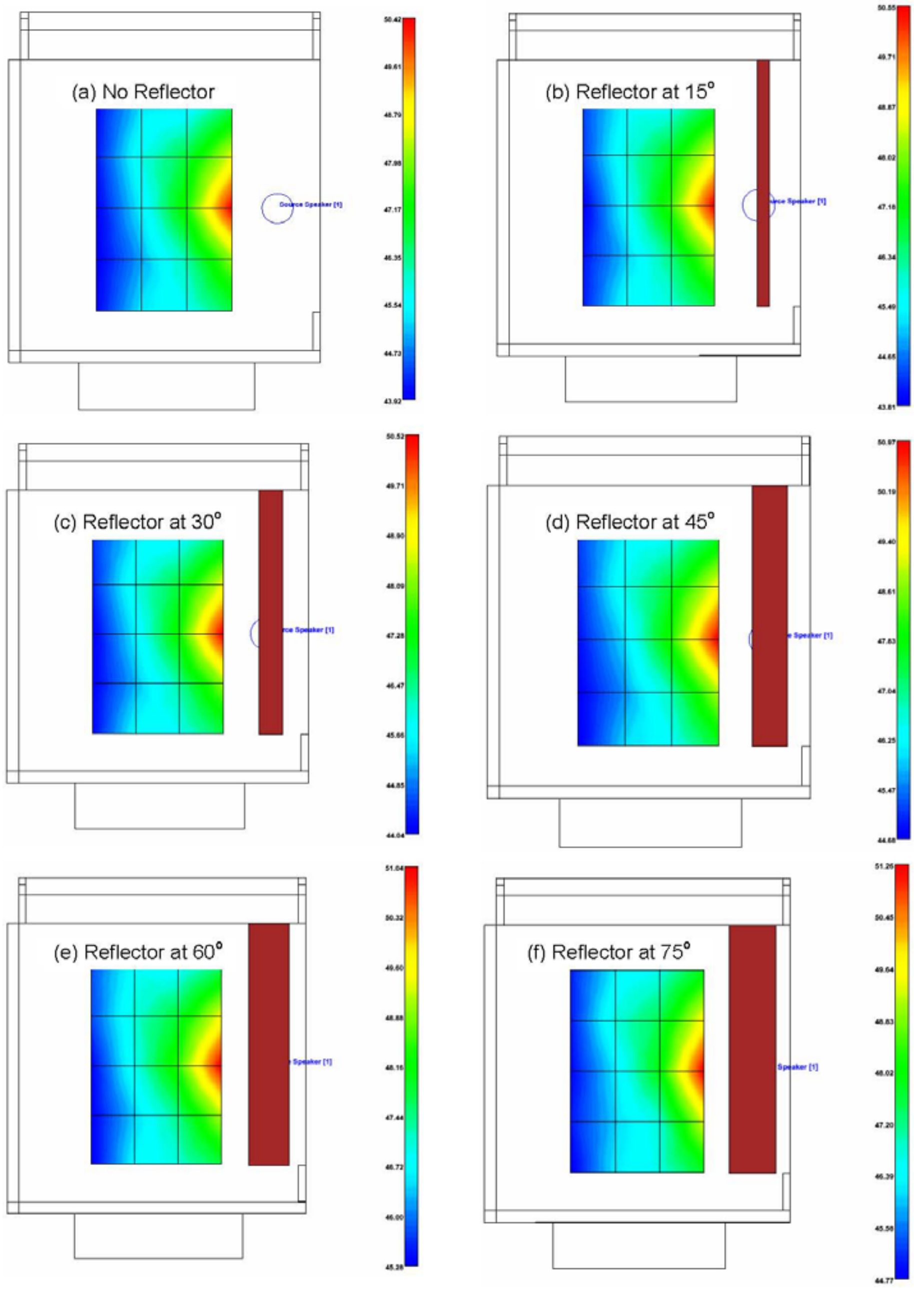


Figure 4.6 – Simulated voice levels in dB at 1k Hz inside the mock-up classroom (with absorptive ceiling and back wall panel) for reflector at different oblique angles.

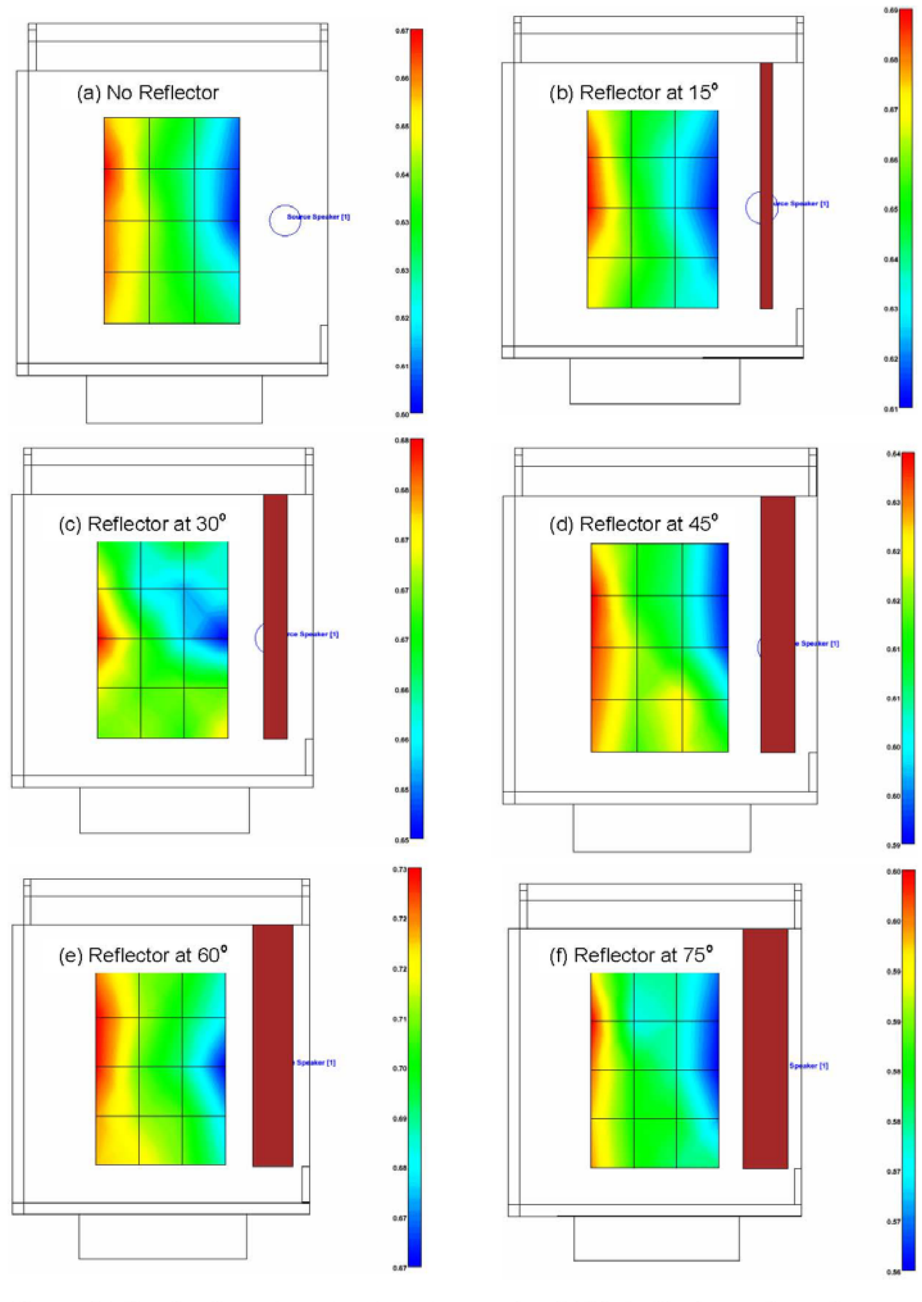


Figure 4.7 – Simulated reverberation times in second at 1k Hz inside the mock-up classroom (with absorptive ceiling and back wall panel) for reflector at different oblique angles.

Chapter 5
Subjective speech intelligibility tests in the selected classrooms

5.1 Introduction

In the previous chapters, acoustical measurements in the selected classrooms, simulation and prediction by ray-based computerized program have been carried out for study on the acoustics configuration and performances of the selected classrooms with different acoustics treatment. Though the speech transmission can be predicted by an objective index – speech transmission index, which has been widely used in recent years, the quality of speech perception is subjective on different person since they vary in age, education background, status of health especially their hearing system, mother tongue language, etc.

The best way to access "real" situation on quality of speech transmission is by human perception. The simplest method of assessment is to conduct a subjective listening test which is the most commonly used. Different types of subjective test have been developed such as the Harvard Intelligibility Test of Phonetically Balanced Words [1], the Fairbanks Rhyme Test [2], the Modified Rhyme Test [3], and the Harvard Test Sentences [1]. Numerous studies on subjective tests in different languages have been conducted through different types of materials [4 – 12] mostly in English. In the past decade, Mandarin and other languages have been studied as a consequence of its considerably boost due to increasing need of mother tongue as educational media have been concerned.

The subjective assessment on quality of speech communication reflects the “real” experiences of human perception; however, it is more difficult and time consuming than objective field measurements. The structure of the tested building requires a great deal money for retrofitting if poor speech transmission is found after assessment. It is vital to have estimation on design stage of building, but not retrofitting of built structure. The relationship between acoustical parameters and speech transmission index have been investigated in certain researches [13 – 17], however, it is not enough to identify the relationship with subjective listening tests’ scores and objective acoustical parameters. The importance of establishing a relationship between acoustical measures and subjective listening test is evident.

It is noticeable that age’s difference leads to different interpretation on the speech materials when they listened. The speech perception is comparatively better for young adults than children due to development of hearing systems [19 – 20]. Based on all kind of tests’ materials, using monosyllabic PB words for listening tests is comparatively reliable because the result indicates more accurately the phonemes that the subject actually heard [1]. In addition, the list of the PB words has a distinguished characteristic: the choice of words reflects the relative frequencies of occurrence of phonemes and their distribution in normal speech used daily. The test results for PB monosyllabic words are comparatively more reliable for speech intelligibility assessment than tests using short words or sentences as they are very sensitive to signal-to-noise ratio [17]. As a result, a more accurate and precise interpretation of the speech intelligibility in a room could be better assessed by articulation scores based on the former material. Hence, PB monosyllabic word lists is used for conducting subjective speech intelligibility tests.

5.2 Material for subjective speech intelligibility tests

The lists of phonetically balanced monosyllabic words for subjective speech intelligibility tests in English and Mandarin are taken from ANSI S3.2-1989 [17] and GB/T 15508-1995 [18] respectively. The word lists in Cantonese have been developed in previous chapter with lists given in Appendix in Chapter 2. There are 20 lists in English with 50 mono-syllabic words in each list. For Mandarin and Cantonese, both languages have ten lists in which 75 characters are selected in each list. It is noticeable that a mono-syllabic word in English is analogous to a Chinese character. The subjective speech intelligibility tests are conducted according to a procedure where the PB monosyllabic words in each list will be presented to a group of students in random order.

5.3 Methodology of subjective speech intelligibility tests

A series of subjective speech intelligibility tests were conducted in the selected classrooms. Two classrooms from a school which were not listed in the previous chapter were selected. There are 5 classrooms from a primary school, a secondary school and a university selected for the test. All classrooms were built in accordance with the Millennium 2000 design except the one in the university. However, their geometrical characteristics and mechanical ventilation equipment inside the classrooms are similar. There are 4 split type air-conditioners and 2 fresh air pre-conditioners in each classroom. Different acoustics treatments are covered in these 5 selected classrooms, including acoustic ceiling, acoustic back wall panel, and carpet. Table 5.1 shows the details of the selected classrooms.

The “subjects” for the tests are required to sit in a selected area inside the classroom such that the boundaries of the zone were about 2 m from the front wall, about 2 m from the back walls, and about 1 m from the windows and doors. Twenty points are evenly selected in the ‘confined’ zone and twelve of them are occupied by the students to conduct the subjective speech intelligibility tests. Figure 5.1 shows a sketch of the layout plan of the experimental set-up.

Prior measurements are then conducted to determine the acoustical parameters at the selected seating locations under comparable background noise conditions without the presence of the subjects. The difference in background noise levels inside the classroom are obtained by alternately 1) turning on or off the indoor mechanical equipment, and 2) opening or closing the windows and doors in various combinations. In addition, a Tannoy loudspeaker producing white noise is placed at the back of the classroom to provide extra background noise levels inside the classroom during the prior acoustical measurements and the subsequent subjective tests. The free field noise level produced by the Tannoy speaker are calibrated at an anechoic chamber with the level set at 60 dB(A), 65 dB(A) and 70 dB(A) measured at 1 m from the speaker. In the measurement of acoustical parameters, the Maximum-length Sequence (MLS) is used as the source signal. The MLS signals is generated by a 01-dB Stell Symphony that is controlled by a notebook computer. The output signal is connected to a B&K 2716 power amplifier and is subsequently broadcasted by an omni-directional sound source (B&K 4296). Same as the extra background noise, the MLS signals is calibrated at an anechoic chamber with the level of 75 dB(A) for configuration in field measurement, except for reverberation time measurement, in the level of 120 dB(A)..

In the subjective tests, the speech is broadcasted by the B&K 4296 omni-directional sound source, where the input speech signals are provided by a notebook computer. The speech for the test is calibrated at an anechoic chamber to maintain equivalent to a sound pressure level of 75 dB(A). A masking background noise level of 60 dB(A), 65 dB(A) and 70 dB(A) (measured at 1 m in the anechoic chamber) is provided at the back of the classroom. White noise is chosen as the masking noise which is generated by a B&K Sine Random Generator Type 1027. The masking noise is processed via the Behringer Eurorack MX 2004A mixer, Ultra-curve Pro DSP 8024 analyser and the Crest Audio LA-1201 power amplifier. The white noise is then generated by a Tannoy Superdual T300 loudspeaker.

The noise levels outside the classroom and the acoustic signals/voice levels inside the classroom are taken respectively by ½-inch pre-polarised free-field microphones (B&K 4189) with wind screens and ½-inch pre-polarised diffuse-field microphones (B&K 4942). The microphones are coupled with DeltaTron preamplifiers

(B&K 2671) that are then connected to a Sony SIR-1000I 16-channels digital data recorder. The recorded raw data are down-loaded and analysed by the 01dB-Stell Signal/Frequency Analyser and Building Acoustics Analyser. The details of instruments are listed in Table 5.2. Picture 5.1 and Picture 5.2 show the experimental setup for the acoustical measurements in the selected classrooms.

During the measurements, the omni-directional sound source and the loudspeaker were set at 1.5 m and 1.2 m above the floor level respectively, which is about the mouth level of an adult and an ear level of the students when seated. The four diffuse-field microphones inside the classroom are located at 1.2 m above the floor level. The free-field microphone on the corridor is set at 1.2 m above the floor level but the other one is protruded at 1 m from the window of the neighbour classroom. All instruments are calibrated before and after all tests as a precautionary measure for ensuring no noticeable drifts in all electronic equipments.

The age of the ‘subjects’ varies between 5 and 22 years old. The study on the speech intelligibility is based on the results from the measurements of acoustical parameters inside a classroom with different background noise levels and subjective experiments for a group of listeners in the same classroom under the same background noise levels for the acoustical measurements. There are 5 different conditions for each group of listeners: a). all windows and doors are opened for natural ventilation, mechanical equipment inside the classroom is switched off, and speech in 75 dB(A) without background masking noise; b). all windows and doors are closed, mechanical equipment inside the classroom is switched on for ventilation, and speech in 75 dB(A) without background masking noise; c). all windows and doors are closed, mechanical equipment inside the classroom is switched off, and speech in 75 dB(A) with background masking noise in 60 dB(A); d). all windows and doors are close, mechanical equipment inside the classroom are switched off, and speech in 75 dB(A) with background masking noise in 65 dB(A); and e). all windows and doors are close, mechanical equipment inside the classroom is switched off, and speech in 75 dB(A) with background masking noise in 70 dB(A).

For each subjective speech intelligibility test, twelve subjects are seated in the pre-selected locations in the classrooms. They are then asked to listen to a pre-recorded speech of the chosen phonetically balanced (PB) monosyllabic words. These PB words are picked from the respective lists of PB Words for Cantonese, Mandarin and English. The PB words of different languages are spoken by their respective native speakers and are digitally taped by a Sony digital data recorder in the anechoic chamber. Different sets of speeches of PB words in English, Mandarin and Cantonese are played in turn under different conditions. Picture 5.3 shows a snapshot of the process when a group of the subjects took the speech intelligibility tests.

During the speech intelligibility tests, diffuse-field microphones are placed at the seats of the selected zone inside the classroom. In addition, a B&K Head and Torso Simulator with a binaural transducer is placed near the centre of the selected zone. The diffuse-field microphones and the Head and Torso Simulator is used to monitor the background noise levels and signal levels in the classrooms, and recorded the signal for analysis of speech intelligibility index. The signal-to-noise ratio (SNR) is the most important acoustical parameter measured and correlated with the scores of subjective speech intelligibility tests to deduce the effect of the acoustical parameters on the speech intelligibility.

5.4 Results on acoustics measurements and subjective tests

The speech intelligibility is assessed through the use of phonetically balanced word scores. They are the percentage of words/characters which are correctly identified by the students in the subjective tests. The effect of signal-to-noise ratio on the speech intelligibility can then be deduced by plotting the PB word score against it for different age groups.

The effect of outdoor noise and noise from mechanical ventilation system is significant to signal-to-noise ratio. As shown in condition (a) and condition (b) of Figure 5.2 to Figure 5.6, the signal-to-noise ratios are observably low nearby the doors, the windows and under split type air conditioners. Although the effect is trivial, the deviation among the whole classrooms are within about 4 dB(A). When the acoustics condition is stringent, which is condition (e), the signal-to-noise ratio is evenly distributed that without large deviation in the classroom with acoustics ceiling. It is implied that acoustics ceiling is effective to maintain an evenly distribution of sound energy inside a classroom. With carpet inside the classroom, there is a better effect on “balance” the sound distribution than with acoustic back wall panel. This finding reinforces the discussion in the previous chapters that acoustics ceiling and carpet have better improvement on speech intelligibility due to its control function on evenly signal-to-noise ratio inside the classroom. Their performance are better than acoustic back wall panel due to its comparatively large in area, and its absorption characteristics are sensitive to frequency bands of human speech.

Figure 5.7 shows the reverberation time of the selected classrooms. All the selected classrooms have unsatisfactory long in reverberation. The classroom P6-B and U1-A are the worst with long reverberation. Some seats have reverberation time almost 0.7 seconds or higher. As previous chapter mentioned, it is unacceptably high as compared with all national and international standards or guidelines for classroom acoustic. Most of the worst positions are located in the side nearby the wall, and middle rear part in the classrooms. It is possibly due to the interior structure of the classrooms. In the G2000 design blueprint, all classrooms have crossbeam cross over in the middle at ceiling. It may create an obstacle that diffusion of signal occurs.

Figure 5.8 and Figure 5.9 illustrate that the perception of the sound signals by students are highly dependent on the level of the signal-to-noise ratios. The PB word score is one reliable indicator to reflect the intelligibility or perception of speech in classrooms. The results showed a clear trend: the higher the signal-to-noise level, the higher PB word scores can be achieved for young children. Furthermore, there is a regular pattern that higher signal-to-noise levels are required for younger children in order to achieve the same PB word scores. The figures show that at least the signal-to-noise ratio should reach a level of at least 11 dB(A) in order to achieve a perfect speech perception of 100% PB word score. In general, a more stringent condition is required for better speech perception inside classrooms for younger children.

For the results, we can see more convincing correlations between speech intelligibility and the measured acoustical parameters for the various age groups of young adults. Based on comparison of different age groups, it is remarkable that the young adults achieved stable in PB word scores earlier in smaller signal-to-noise ratio. The main reasons for this phenomenon being the development of hearing systems is matured for young adults, but not yet well developed in children. Besides, the knowledge and experiences with age growth may help their ability in guessing the words they heard.

Through comparison on different languages, the PB word scores in Cantonese which is shown in Figure 5.8 demonstrates an earlier stability in lower signal-to-noise ratio than in English. Besides the fact that Cantonese is the mother tongue of the students, the tonal characteristics are also taken into account on their effect to speech transmission. Cantonese and Mandarin are tonal languages from same the language family. The presence of tones helps listeners to predict or guess the words they listen, as opposed to English which is a non-tonal language. It is hard to guess without tones' effect for phonetically balanced monosyllabic words test.

5.4 Conclusions

From the figures shown, it is suggested that the signal-to-noise ratio shows a close correlation with PB word score. It is possible to conclude that the higher the signal-to-noise ratio, the better the PB word scores can be achieved by the subjects of different age groups. It also implies that the subjective speech intelligibility test with PB word scores is an effective mean to measure the real situation of acoustics quality, or speech quality inside the classrooms. It is clearly stated that the earlier age of students, the stringent conditions in signal-to-noise ratio have to be considered. The signal-to-noise ratio should be higher than 11 dB(A) for over 70% of speech perception by PB word scores calculation.

Besides the relationship between PB word scores and signal-to-noise ratio, it is ponderable to review the situation and material used in the current selected classrooms. All the selected classrooms are unacceptably high in reverberation. Although acoustics ceiling and carpet are comparatively effective to maintain an even distribution of signal-to-noise ratio and decrease the reverberation by their absorption characteristics, the reverberation inside the classrooms is far from a satisfactory level. Long reverberation time interfered the continuous signal of the speech as masking signal, and cocktail party effect like echo is easily created in high percentage in hard wall inside the classrooms, which affect the speech intelligibility.

References

1. EGAN, J. P., “Articulation Testing Methods,” Laryngoscope 58, pp. 955-991 1948.
2. FAIRBANKS, G., “Test of Phonemic Differentiation: The Rhyme Test,” J. Acoust. Soc. Am. 30, pp. 596-600, 1958.
3. HOUSE, A. S., WILLIAMS, C. E., HECKER, M. H. L. and KRYTER, K. D., “Articulation Testing Methods: Consonantal Differentiation with a Closed-Response Set,” J. Acoust. Soc. Am. 37, pp. 158-166, 1965.
4. STEENEKEN, H. J. M. and HOUTGAST, T., “A physical method for measuring speed-transmission quality,” J. Acoust. Soc. Am. **67**, pp. 318-326, 1980.
5. ANDERSON, B. W. and KALB, J. T., “English verification of the STI method for estimating speech intelligibility of a communications Channel,” J. Acoust. Soc. Am. **81**, pp. 1982 – 1985, 1987.
6. KANG, J., “Comparison of speech intelligibility between English and Chinese,” J. Acoust. Soc. Am., **103**, pp. 1213 – 1216, 1998.

7. HOUTGAST, T. and STEENEKEN, H. J. M., "A multi-language evaluation of the RASTI-method for estimating speech intelligibility in auditoria," Acustica **54**, pp. 185 – 199, 1984.

8. ZHANG, J. L., "Speech" in "Handbook of Acoustics", edited by MAA, D.Y. and SHEN, H., Science Press, Beijing, Ch. 19, pp. 404-435, 1987.

9. BERANEK, L. L., "Acoustic Measurements", Wiley, New York, Ch. 13, pp. 625-634, Ch. 17, pp. 761-792, 1949.

10. HOUTGAST, T. and STEENEKEN, H. J. M., "A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria", J. Acoust. Soc. Am **77**, pp. 1069-1077, 1985.

11. HOUTGAST, T., "The effect of ambient noise on speech intelligibility in classrooms", Applied Acoustics **14**, pp. 15-25, 1981.

12. BRADLEY, J. S., "Speech intelligibility studies in classroom", J. Acoust. Soc. Am. **80**, pp. 846-854, 1986.

13. KOOTWIJK, P. A. A., “The speech intelligibility of the public address systems at 14 Dutch railway stations,” J. Sound Vib. 193, pp. 433 – 434, 1996.

14. IEC Standard 60268-16, 2nd Edition. Sound System equipment – Part 16: Objective rating of speech intelligibility by speech transmission index, 1998.

15. STEENEKEN, H. J. M., and HOUTGAST, T., “Mutual dependence of the octave band weights in predicting speech intelligibility,” Speech Comm. 28, pp. 109 – 123, 1999.

16. STEENEKEN, H. J. M., and HOUTGAST, T., “Phoneme-group specific octave-bands weights in predicting speech intelligibility,” Speech Comm. 38, pp. 399 – 411, 2002.

17. ANSI S3.2, American National Standard: Method for measuring the intelligibility of speech over communication systems, American National Standards Institute, New York, 1989.

18. GB/T 15508, National Standard of Peoples Republic of China: Acoustics – Speech articulation testing method, National Technical Supervision Bureau, 1995.

19. MOORE, D. R., “Auditory development and the role of experience”, British Medical Bulletin **63**, pp. 171-181, 2002.

20. BAILEY, P. J. and SNOWLING, M. J., “Auditory processing and the development of language and literacy”, British Medical Bulletin 63, pp. 135-146, 2002.

Tables

School Code	**Classroom Code**	**Design**	**Measured Volume / m^3**	**Mechanical Ventilation System**	**Acoustic Treatment**
P6	P1 – A	standard	211	2 A/C & 2 FAP	acoustic back wall panel
	P1 – B				-
S5	S5 – A	standard	211	2 A/C & 2 FAP	acoustic ceiling
	S5 – B				
U1	U1 – A	non-standard	239	2 A/C & 2 FAP	carpet

Note:

[1] P indicates primary school, S indicates secondary school and U indicates university.

[2] A/C represents split type air-conditioner, FAP represents fresh air pre-conditioner.

Table 5.1 – Geometrical and acoustical details of the classrooms for subjective speech intelligibility tests.

Manufacturer	Type	Qty.
Sony	16 Channels Sensor Input High Speed Digital Data Recorder Type SIR-1000I	1
Bruel and Kjaer	½” Pre-polarised Free-Field Microphone Type 4189 with wind screen	2
	½” Pre-polarised Diffuse-Field Microphone Type 4942	3
	DeltaTron Preamplifier Type 2671	5
	Noise Calibrator Type 4231	1
01 dB - Stell	Signal/Frequency Analyser: dBFA32	
	Building Acoustics Analyser: dBBATI32	

Table 5.2 – Instruments and processing software used for acoustics measurements in the classrooms.

Figures

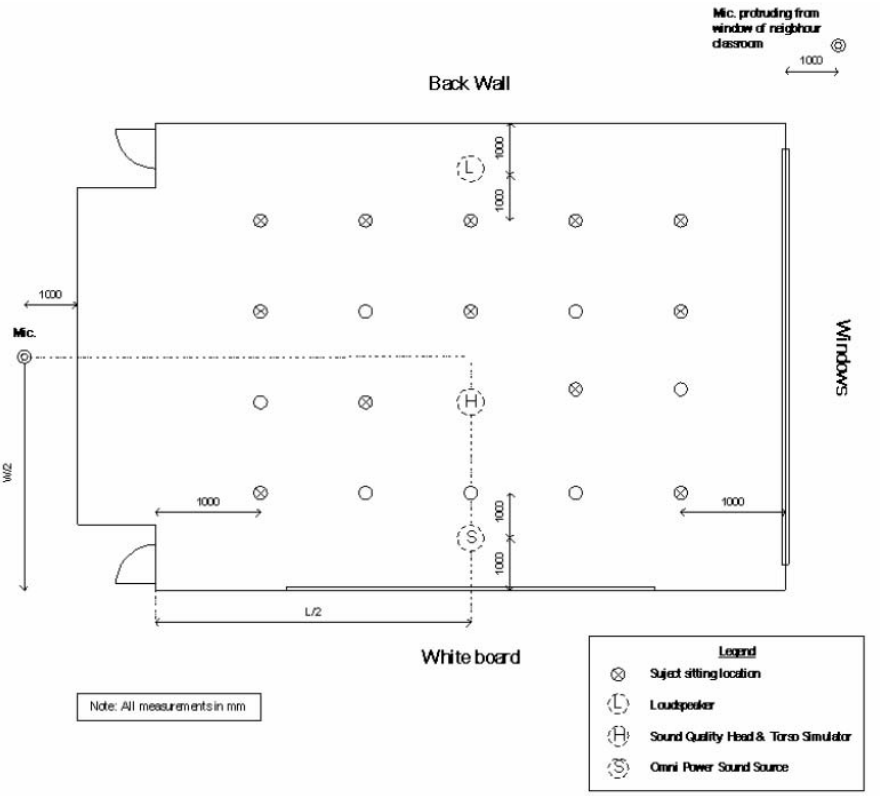
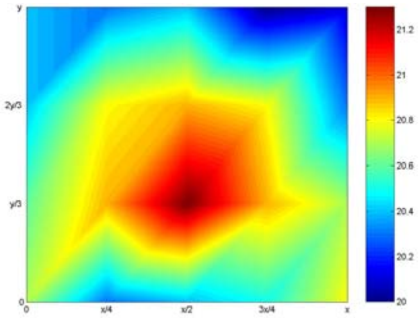
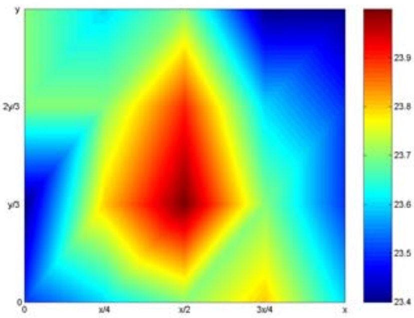


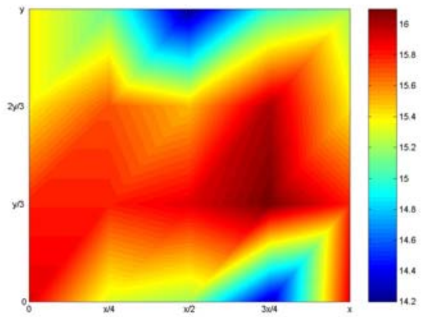
Figure 5.1 – Details of experimental setup in the selected classrooms for subjective speech intelligibility tests.



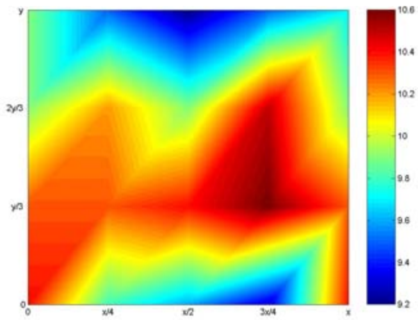
(a)



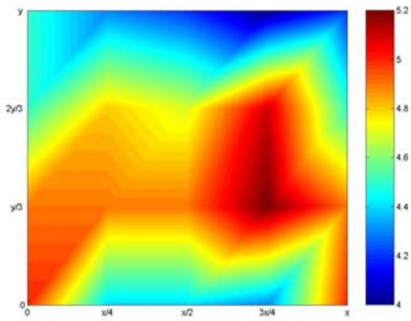
(b)



(c)



(d)

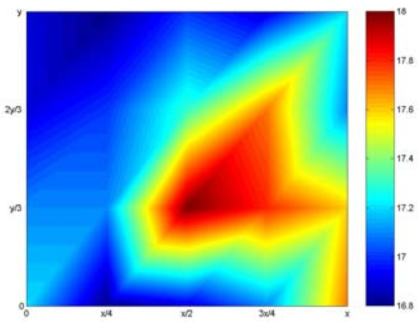


(e)

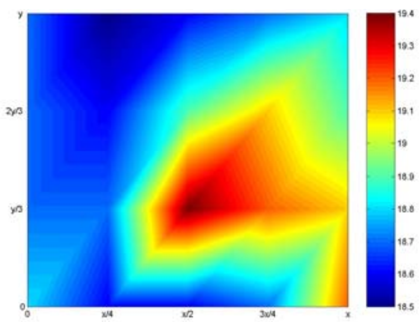
Figure 5.2 – Signal-to-noise ratio of the classroom P6-A under conditions:

(a) windows open, mechanical ventilation off, signal in 75 dB(A);
(b) windows closed, mechanical ventilation on, signal in 75 dB(A);
(c) windows closed, mechanical ventilation off, signal in 75 dB(A) with background masking noise in 60 dB(A);
(d) windows closed, mechanical ventilation off, signal in 75 dB(A) with background masking noise in 65 dB(A);
(e) windows closed, mechanical ventilation off, signal in 75 dB(A)

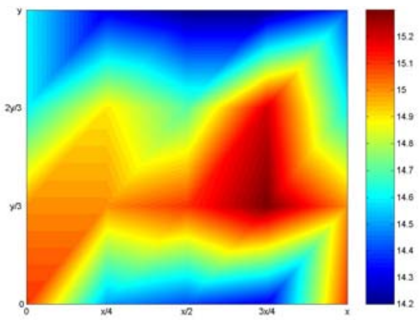
with background masking noise in 70 dB(A).



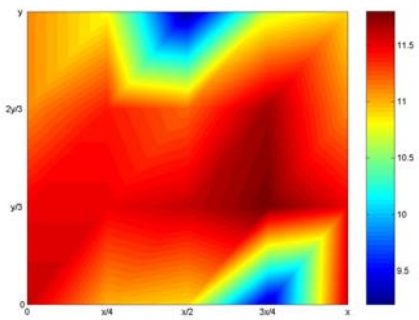
(a)



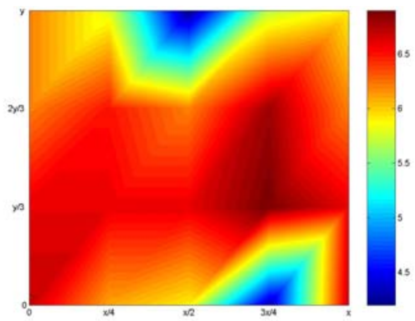
(b)



(c)



(d)

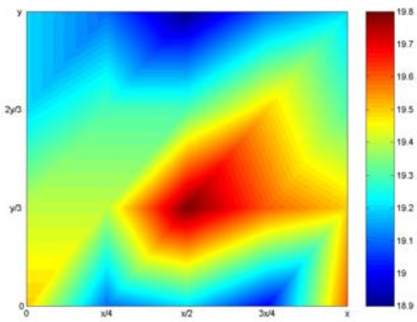


(e)

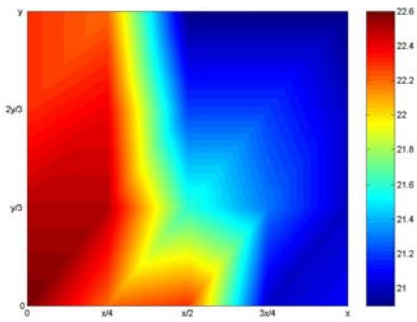
Figure 5.3 – Signal-to-noise ratio of the classroom P6-B under conditions:

(a) windows open, mechanical ventilation off, signal in 75 dB(A);
(b) windows closed, mechanical ventilation on, signal in 75 dB(A);
(c) windows closed, mechanical ventilation off, signal in 75 dB(A) with background masking noise in 60 dB(A);
(d) windows closed, mechanical ventilation off, signal in 75 dB(A) with background masking noise in 65 dB(A);
(e) windows closed, mechanical ventilation off, signal in 75 dB(A)

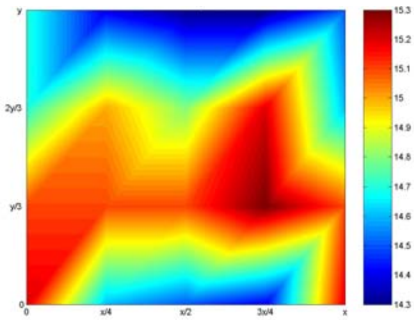
with background masking noise in 70 dB(A).



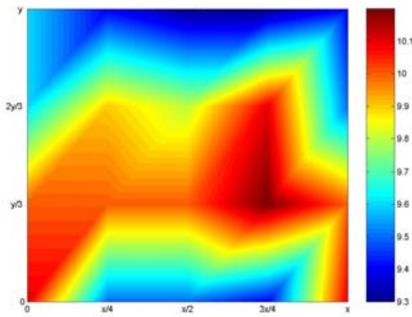
(a)



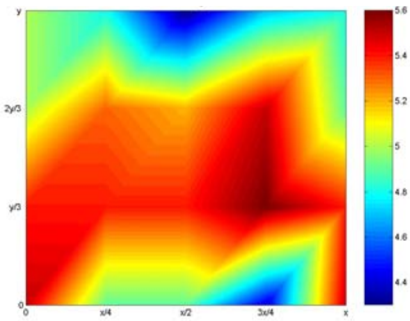
(b)



(c)



(d)

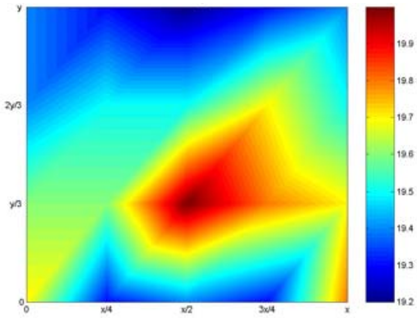


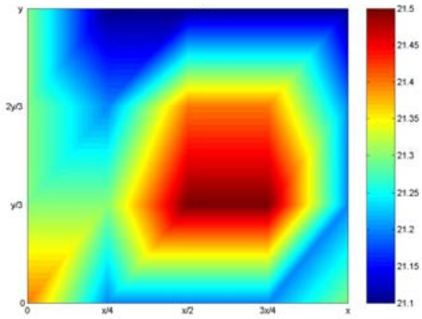
(e)

Figure 5.4 – Signal-to-noise ratio of the classroom S5-A under conditions:

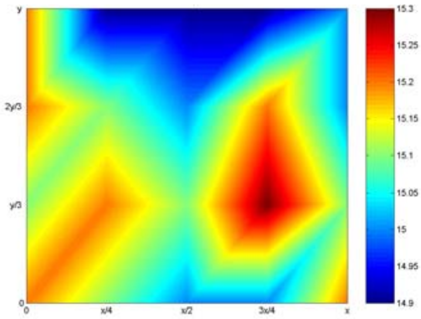
(a) windows open, mechanical ventilation off, signal in 75 dB(A);
(b) windows closed, mechanical ventilation on, signal in 75 dB(A);
(c) windows closed, mechanical ventilation off, signal in 75 dB(A) with background masking noise in 60 dB(A);
(d) windows closed, mechanical ventilation off, signal in 75 dB(A) with background masking noise in 65 dB(A);
(e) windows closed, mechanical ventilation off, signal in 75 dB(A)

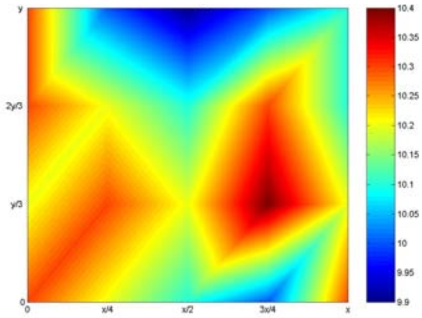
with background masking noise in 70 dB(A).



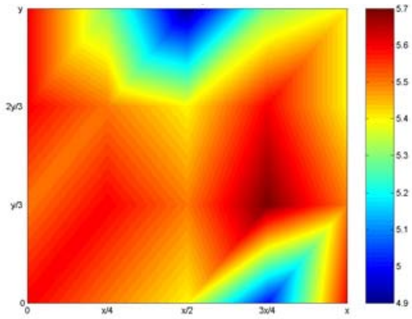


(a) (b)





(c) (d)

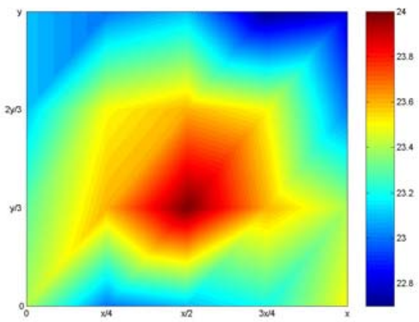


(e)

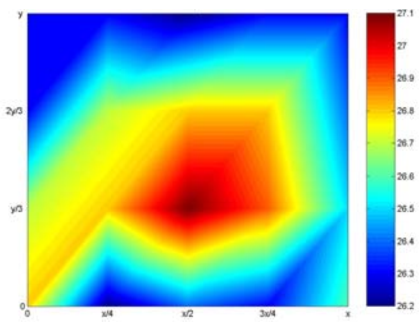
Figure 5.5 – Signal-to-noise ratio of the classroom S5-B under conditions:

(a) windows open, mechanical ventilation off, signal in 75 dB(A);
(b) windows closed, mechanical ventilation on, signal in 75 dB(A);
(c) windows closed, mechanical ventilation off, signal in 75 dB(A) with background masking noise in 60 dB(A);
(d) windows closed, mechanical ventilation off, signal in 75 dB(A) with background masking noise in 65 dB(A);
(e) windows closed, mechanical ventilation off, signal in 75 dB(A)

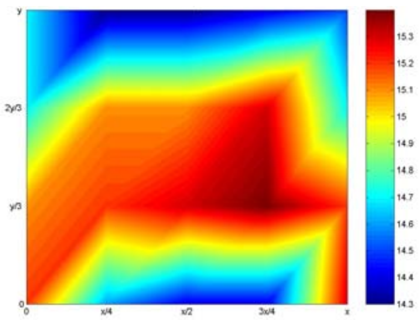
with background masking noise in 70 dB(A).



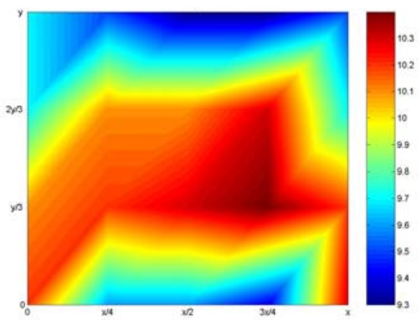
(a)



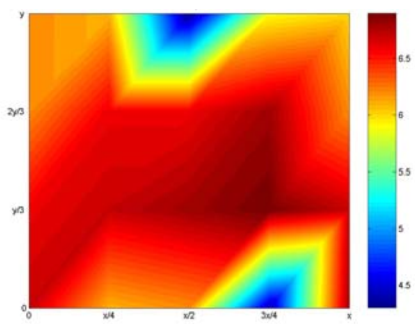
(b)



(c)



(d)

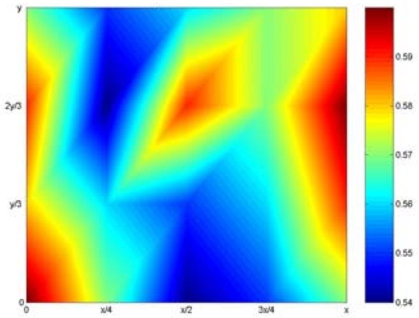


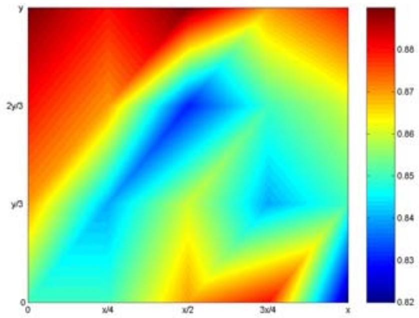
(e)

Figure 5.6 – Signal-to-noise ratio of the classroom U1-A under conditions:

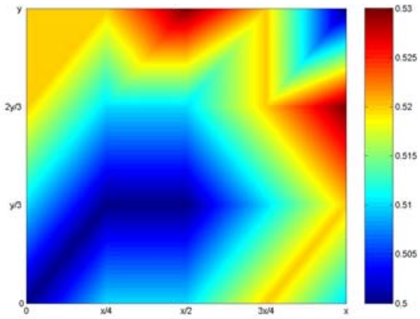
(a) windows open, mechanical ventilation off, signal in 75 dB(A);
(b) windows closed, mechanical ventilation on, signal in 75 dB(A);
(c) windows closed, mechanical ventilation off, signal in 75 dB(A) with background masking noise in 60 dB(A);
(d) windows closed, mechanical ventilation off, signal in 75 dB(A) with background masking noise in 65 dB(A);
(e) windows closed, mechanical ventilation off, signal in 75 dB(A)

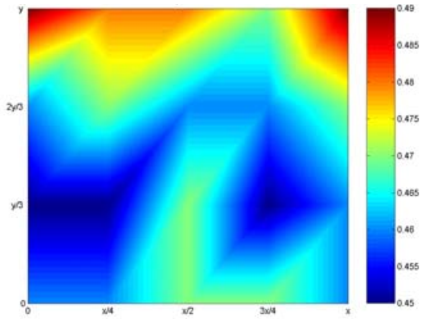
with background masking noise in 70 dB(A).



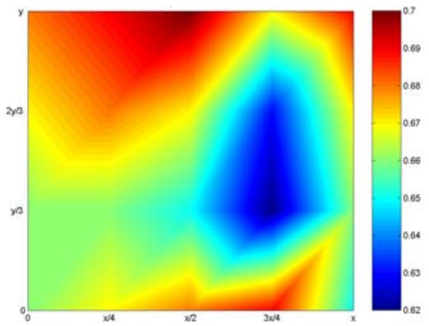


(a) (b)





(c) (d)



(e)

Figure 5.7 – Reverberation time of different classrooms:

(a) P6-A; (b) P6-B; (c) S5-A; (d) S5-B; (e) U1-A.

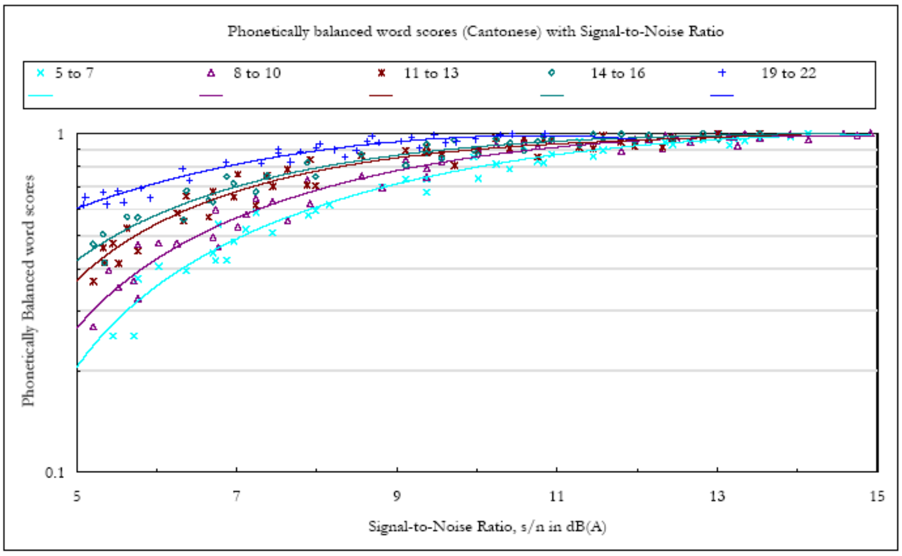


Figure 5.8 – Phonetically balanced word scores of the subjects in Cantonese.

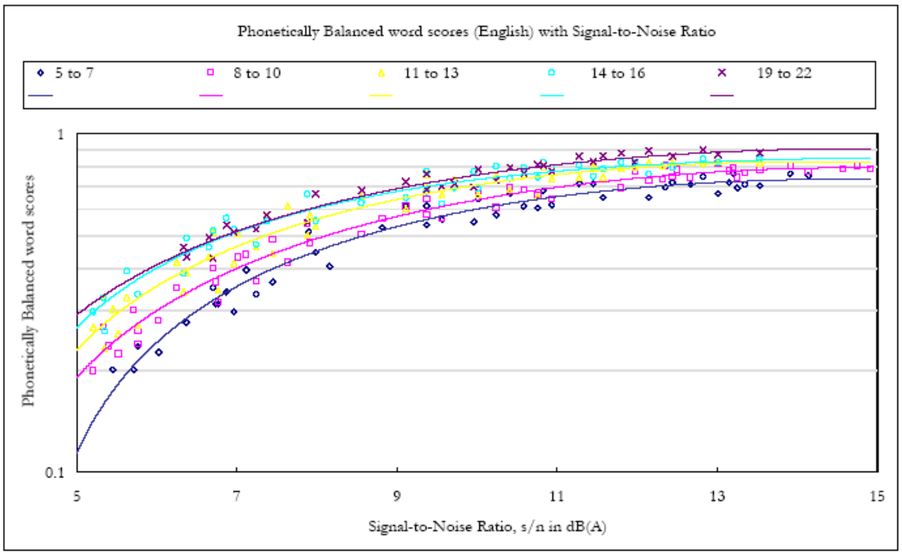


Figure 5.9 – Phonetically balanced word scores of the subjects in English.

Pictures

Picture 5.1 – The signal and noise generation system used for the acoustical measurements and subjective speech intelligibility tests.

Picture 5.2 – Experimental setup for acoustics measurements in the selected classrooms before subjective speech intelligibility tests.

Picture 5.3 – Students sitting for the subjective speech intelligibility tests in the selected classrooms.

Chapter 6
Data modelling on effect of acoustics parameters to speech intelligibility

6.1 Introduction

Nowadays, study and application of research data processing and analysis becomes complex in the movement of technology development. The huge database and various relationships between the data lead a long computation time to find out a solution. In recent years, application of computerize program such as neural network is widely applied as a tool to research in cognitive science and engineering. Neural network modelling has numerous advantages than tradition methods on experimental data modelling. It's architecture based on learning and training abilities through the experiences of the data. The network is then able to "simulate" and "memorize" the input variables and output variables of any complex functional relationship, to deal with a variety of vague and non-linear data through "simulation" and "estimation" to predict the results.

Speech intelligibility is a subjective human perception experience to quantify the quality of speech transmission in a place. Classrooms, theatres, legal courts, hospitals, stations of public transportation, etc. are the most influential places to speech perception. With poor speech communication, misunderstanding between people will increase; accidents will happen more easily; workforce will become inefficient; mental illness will be increase in number due to annoyance; and education of new generation will be retrogressive. However, the assessments of speech intelligibility involve enormous data from subjective listening tests which includes useful data and numerous human errors. Applications of neural network model to study the speech intelligibility have been researched in the past decades. Li and Cox used neural network model to quantify the speech intelligibility from transmitted speech signals and various speech transmission index [1 – 2]. Xu founds the neural network approach is accurate to predict the speech intelligibility and security thresholds with functional relationship between signal-to-noise ratio and octave frequencies [3]. Li and Cox further investigated that neural network model is accurate to predict speech transmission index from running speech simulation [4]. Application of neural network model on medical related aspect is noteworthy in many researches. Metz research findings on its application to estimate the speech intelligibility of hearing impaired speakers from acoustics parameters identified the application of neural network modelling is useful on study related to human response [5]. Alexandre et al. developed an automatic sound classification system for hearing aids in different environment simulation [6]. Kocinski researched on improving speech intelligibility using speech and speech-shaped noise signal which aim on application of hearing aids [7].

This chapter will first give generalized regression neural network model to the variables surveyed in previous chapters, and conduct a simulation by MATLAB on algorithm and model to study of phonetically balanced word scores, signal-to-noise ratio, reverberation time and speech transmission index.

6.2 Introduction of GRNN

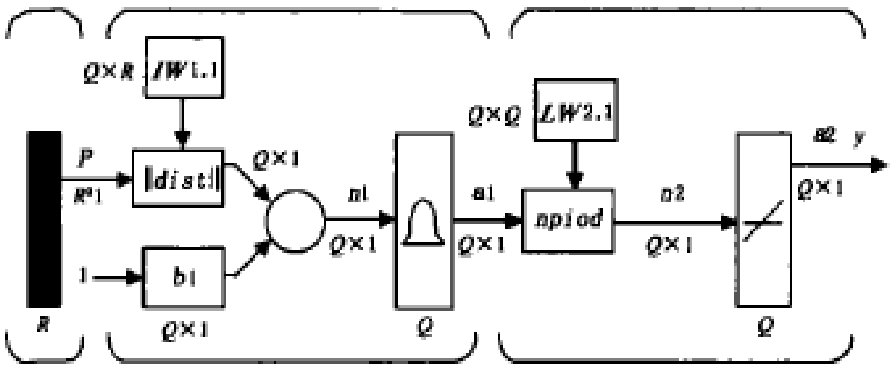
6.2.1 Selection of network model

Feed forward back propagation neural network is one of the networks widely used; however, there have problems that plagued the algorithm: 1) convergence rate; 2) insufficient points in certain area; 3) network paralysis; 4) stability; and 5) step-size. These factors bring out unreliability and inaccuracy on the simulation results. Refer to a trial test on using this network to simulate the experiment data, the result is unsatisfactory and its learning ability is not reliable due to the above factors.

Generalized regression neural network (GRNN) is selected as model for the current study. GRNN requires a simple parameter that a long cycle on training process can be neglected. The calculation speed and its advantages of stability on the result makes its extensive application on non-linear function approximation, data classification, pattern recognition, system modelling and other complex area.

6.2.2 Algorithm of GRNN

Generalized regression neural network, or probabilistic neural network is developed by Specht in 1990. The structure and algorithm is radial basis network approach that suitable for functional approximation. It forms by input layer, radial basis hidden layer and linear output layer where P is input vector, R is network input dimension and Q is number of training samples.



The number of units or elements of radical basis hidden layer is the training samples size Q. The weighted function of the layer is Euclidean distance function shown in “||dist||”. Its role is to calculate the network input and hidden layer weighted matrix $IW_{1.1}$ where 1 is the distance between and b1 is the hidden layer thresholds. Symbol “ · ” is the relationship between ||dist|| output and threshold elements b1. The results of the formation are input to n1 and send to transfer function. The hidden layer transfer function is the radial basis function in Gaussian functions, such as

$$R_i(x) = \exp(-\frac{\| x - c_i \|^2}{2\sigma_i^2})$$

The equation shows σ_i decides the shape of i-th of basis function in the hidden layer. The greater the σ_i, the basis function becomes smoother and σ_i is also known as smoothness. The weighted function of network output layer is the standardization of points multiplication function in "nprod". It computes the network vector n^2 by the vector a1 and weighted matrix $LW_{2.1}$. The result provide a linear transfer function a2 = purelin(n2) for computing the network output.

The learning ability of connection weighted functions in GRNN is using feed forward back propagation algorithm since the nodes in the radial basis hidden layer are in Gaussian function. Gaussian function is non-negative non-linear function of partial distribution on radial symmetry attenuation or decline. The partial response will be generated from input signal. The hidden layer nodes will have a larger output when the input signal close to basis function of the central area. It demonstrates the network has ability in local approximation and it is the reason on comparative fast in learning and calculation speed.

Besides, there is only one threshold in GRNN for adjustment that the network is solely depends on learning from the data. This characteristic avoids subjective assumptions to affect the prediction results.

6.3 Neural network model on phonetically balances word scores

6.3.1 Methodology on neural network model setup

To setup a generalized regression neural network modelling, the data have to pre-processing according to the characteristics of the original data to select for training samples which include input data and target output data. There have several methods on data pre-processing, though, the normalization is chosen in this study for pre-processing method that all the input and output data have to be trained and transformed in to range [0, 1]. The input and output data dimension of the network is determined by training samples. The appropriate trial samples are selected with the corresponding data pre-processing. The smoothness is the most important parameters to affect the performance and accuracy of the network. With smaller smoothness, the network is more accurate in approximation on the samples size. With larger smoothness, the network is more accurate in approximation on the data. Therefore, a continuous trial is necessary for selection on best value of smoothness to balance the simulation environment. The network is verified on its accuracy after training and learning. The network is used to simulate the trial data and comparative analysis is carried out between the expected results and the output from the network. Throughout the comparison, the network meets the expectation and requirement.

6.3.2 Neural network model simulation and result

From the floor plan of the selected classrooms for subjective speech intelligibility tests, the appropriate data are selected as training samples of the network. Particular seats like 3, 11 and 19 are ignored for the potential deviation of acoustical characteristics. The objective of the network is to study the relationship and their effect of signal-to-noise ratio (SNR), reverberation time (RT) and speech transmission index (STI) on phonetically balanced word scores (PB). SNR, RT and STI are chosen as input variables P and so they constitutes a three-dimensional vector P. Meanwhile, the corresponding SNR, RT and STI of the selected seats are numbering for input variables in the network.

Since the deviation and range are varies in the data of SNR, RT and STI, direct input of vector P will generate inaccurate simulation and results. Hence, the vector P has to pre-process and normalization is used. Each row of the vector P is assumed as a unit. The data is used to corresponding value of the new vector if the element value of the row vector is 0. Otherwise, each element of row vector is divided by quadratic norm as the corresponding new value of the new vector. After the pre-processing on vector P, it becomes a data in range of [0, 1] for all values.

Although the age difference on PB has been demonstrated in pervious chapter, the trend is similar for each age range under similar relationship with SNR. Thus, the age group 8 to 10 years old are selected for data modelling. The output vector T is therefore in one-dimensional, which is about the PB. The trial data is from the measured database except those selected for training and learning process of the network. Data at seat of dummy head is chosen for trial test on the network. For its input vector P_t (from SNR, RT and STI), the corresponding output vector T_t (from PB) is identified. In the MATLAB neural network toolbox, the GRNN network function is in "newgrnn" which contain 3 parameters: the input vector P, the output vector T and the smoothness.

The smoothness is a very important variable on the accuracy of simulation of the network. To determine the smoothness, repeated trial values and trial networks are used to ensure the accuracy. The smoothness F in 0.1 is firstly used for the determination process. Figure 6.1 shows the result estimated by F = 0.1 with trial data. The curve in green demonstrated the expected variation between the trial data with input P_t and output T_t. Compared with the curve in blue, which is the predicted output T_F with input Pt and smoothness F, it obviously showed a poor estimation in F = 0.1. The standard deviation is about 0.0199 and shown in Figure 6.2. Since the smoothness F ranging from [0.01 to 0.1], several iterations are carried out in different F for the network. With F = 0.05, the estimation is the best which is closed to our trial data with standard deviation about 0.0116, see Figure 6.3 and Figure 6.4.

In order to verify the reliability of the selected smooth fact F = 0.05, data from several seats excepted 3, 11 and 19 are selected as trail for the network. Similar procedures are carried out and it is concluded that F = 0.05 is the best value for the network. The verification result is shown in Figure 6.5 and Figure 6.6.

As Cantonese and Mandarin are tonal languages which opposed to English as it is non-tonal language, previous experimental results from Cantonese and English are chosen for this study. The speech transmission inside a classroom with 'enough' signal-to-noise ratio is reasonably adequate; however, the critical condition for speech transmission is what the study concerned. Accordingly, condition under windows and doors are closed, all mechanical ventilation equipments are closed, speech signal in 75 dB(A) and background masking noise in 70 dB(A) is selected for the network. In order to simulate the variation of SNR, RT and STI for the study, several variations in variables of input vector are created and taken into account to constitute the new input vectors for the network processing. Table 6.1 illustrates the variation in combination of different input vectors.

Compared the results on the effect of PB on different languages, PB in Cantonese and English have similar results on impact of different variation in the acoustics parameters. It is clearly stated that variation in SNR has the greatest effect on PB change. The PB value is increased when SNR is increased. The influence of variation in RT and STI is not significant as in SNR. Table 6.2 illustrated the results of effect of acoustics parameters to the PB in Cantonese and in English.

Several conditions of variation in different classrooms are studied with individual change in acoustics parameters. Four conditions of RT and STI are examined under the same SNR. From the results in Table 6.3 for rear positions, under the same condition of SNR and different values of STI, the impact of RT is insignificant to PB. This is possibly due to the measured data to be in a small range of reverberation time. Besides, the positions selected on the results are at the rear part of the classroom, where reverberation is frequently occurred over the area. The effect of reverberation time to the phonetically balanced word scores is not trivial and noticeable. Nevertheless, the results investigated the effect of reverberation time to phonetically balanced word scores is more considerable than change of speech transmission index. In addition, the classrooms (P6-B and U1-A) without acoustics ceiling and acoustics back wall panel are the worst in speech intelligibility in both languages when reverberation time is shorten. The needs on acoustics treatment as acoustics ceiling and acoustics back wall panel suggested in previous chapters are further reinforced the statement.

To further investigate the relationship and the effect of individual acoustical parameters to speech intelligibility, or phonetically balanced word scores in different age group, analysis on the nature of their relationship found in pervious chapters have been conducted. The value of phonetically balanced word scores is ranged from 0 to 1. With growth of signal-to-noise ratio or speech transmission index, the PB value will eventually tend to 1. From the analysis of their nature in relationship, the empirical relationships between different parameters in individual age group are identified as below. In the empirical equations, x is individual acoustics parameters and y is phonetically balanced word scores.

PB to STI

Age Group	Empirical Formula
5 to 7	y = -0.3683/x + 1.4274
8 to 10	y = -0.2831/x + 1.3707
11 to 13	y = -0.2663/x + 1.3925
14 to 16	y = -0.2586/x + 1.4026
19 to 22	y = -0.2226/x + 1.4234

PB to SNR

Age Group	Empirical Formula
5 to 7	y = 1/(-0.0086 x + 1.4997)
8 to 10	y = 1/(-0.0051 x + 1.2738)
11 to 13	y = 1/(-0.0042 x + 1.1885)
14 to 16	y = 1/(-0.0039 x + 1.1530)
19 to 22	y = 1/(-0.0013 x + 1.0257)

PB to RT

Age Group	Empirical Formula
5 to 7	y = 0.0114/x + 0.7111
8 to 10	y = 0.0087/x + 0.8201
11 to 13	y = 0.0082/x + 0.8746
14 to 16	y = 0.0080/x + 0.8996
19 to 22	y = 0.0040/x + 0.9922

6.4 Conclusions

Through the data modelling and concluded empirical relationship between the acoustics parameters: signal-to-noise ratio SNR, reverberation time RT and speech transmission index STI, and subjective results of speech intelligibility tests, PB word scores, several conclusive remarks are observed.

The phonetically balanced word score is sensitive and has decisive impact of change in signal-to-noise. With larger signal-to-noise ratio, higher phonetically balanced word scores can be achieved that better speech perception is received. Their proportional relationship is clearly shown in the figures. Under the condition of signal-to-noise ratio remain unchanged; the effect of reverberation time is relatively apparent. However, its relationship with speech intelligibility is inverse. With longer reverberation time, the phonetically balanced word score is lower that less speech can be heard it clearly. In contrast, the relationship with speech transmission index is non trivial. Although with higher in speech transmission index, higher phonetically balanced word score is expected to achieve, the influence of speech transmission index is not significant since STI is objective index widely used to quantify the speech transmission. Therefore, speech transmission index should be proportional to the results of subjective speech intelligibility tests in the same position and acoustics conditions.

References

1. LI, F. F. and COX, T. J., "Predicting speech transmission index from speech signals using artificial neural networks", Proceedings of the World Multi-Conference on Systemics, Cybernetics and Informatics SCI, pp. 43-47 2000.

2. LI, F. F. and COX, T. J., "A neural network model for speech intelligibility quantification", Applied Soft Computing **7**, pp. 145-155, 2007.

3. XU, J. F., BRADLEY, J. S. and GOVER, B. N., "An artificial neural network approach for predicting architectural speech security", J. Acoust. Soc. Am. **117**, pp. 1709-1712, 2005.

4. LI, F. F. and COX, T. J., "Speech transmission index from running speech: a neural network approach", J. Acoust. Soc. Am. **113**, pp. 1999-2008, 2003.

5. METZ, D. E. et al., "The use of artificial neural networks to estimate speech intelligibility from acoustics variables: a preliminary analysis", J. Comm. Disord. **25**, pp. 43-53, 1992.

6. ALEXANDRE, E. et al., "Automatic sound classification for improving speech intelligibility in hearing aids using a layered

structure", Lecture Notes in Computer Science **4224**, pp. 306-313, 2006.

7. KOCINSKI, J., "Speech intelligibility improvement using convolutive blind source separation assisted by denoising algorithms", Speech Communication **50**, pp. 29-37, 2008.

Tables

	Combination of variables							
SNR	+	+	+	+	–	–	–	–
RT	–	–	+	+	–	–	+	+
STI	+	–	–	+	–	+	–	+

Table 6.1 – Combination of variables for input vectors.

	Combination of variables							
SNR	+	+	+	+	–	–	–	–
RT	–	–	+	+	–	–	+	+
STI	+	–	–	+	–	+	–	+
Cantonese	+	+	+	+	–	–	–	–
English	+	+	+	+	–	–	–	–

Table 6.2 – Effect of acoustics parameters to PB word scores in Cantonese and in English.

	P6-A	P6-B	S5-A	S5-B	U1-A
Cantonese	–	+	–	–	+
English	–	+	–	–	+

(a) RT increased, STI increased

	P6-A	P6-B	S5-A	S5-B	U1-A
Cantonese	–	+	–	–	+
English	–	–	–	–	+

(b) RT increased, STI decreased

	P6-A	P6-B	S5-A	S5-B	U1-A
Cantonese	+	–	+	+	–
English	+	–	+	+	–

(c) RT decreased, STI increased

	P6-A	P6-B	S5-A	S5-B	U1-A
Cantonese	+	–	+	+	–
English	+	–	+	+	–

(d) RT decreased, STI decreased

Table 6.3 – Effect of acoustics parameters to PB word scores in Cantonese and in English, at rear positions.

Figures

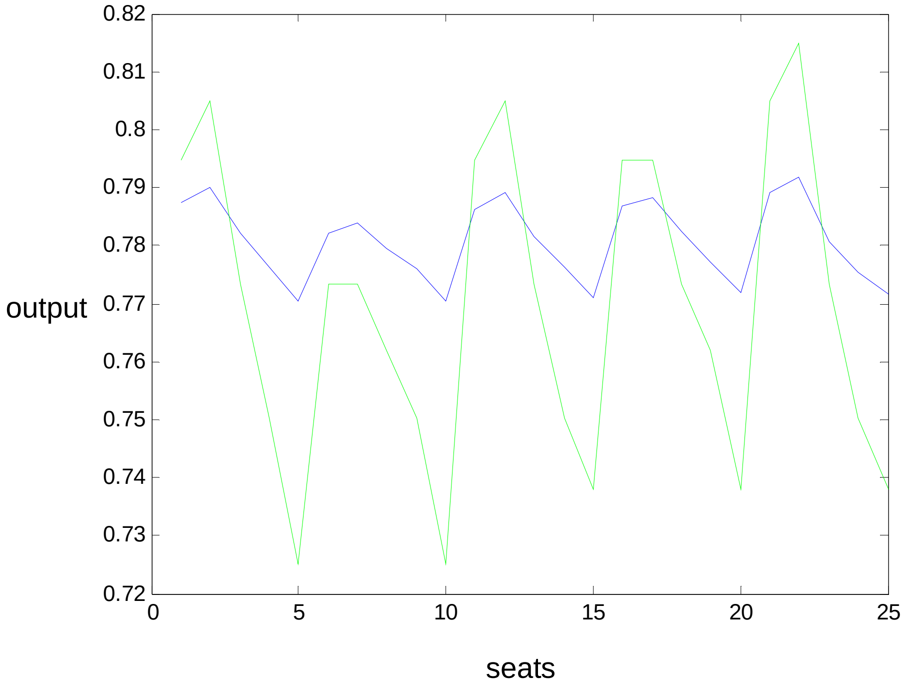


Figure 6.1 – Comparison of simulation results in F = 0.1. (Green – trial data, Blue – predicted data)

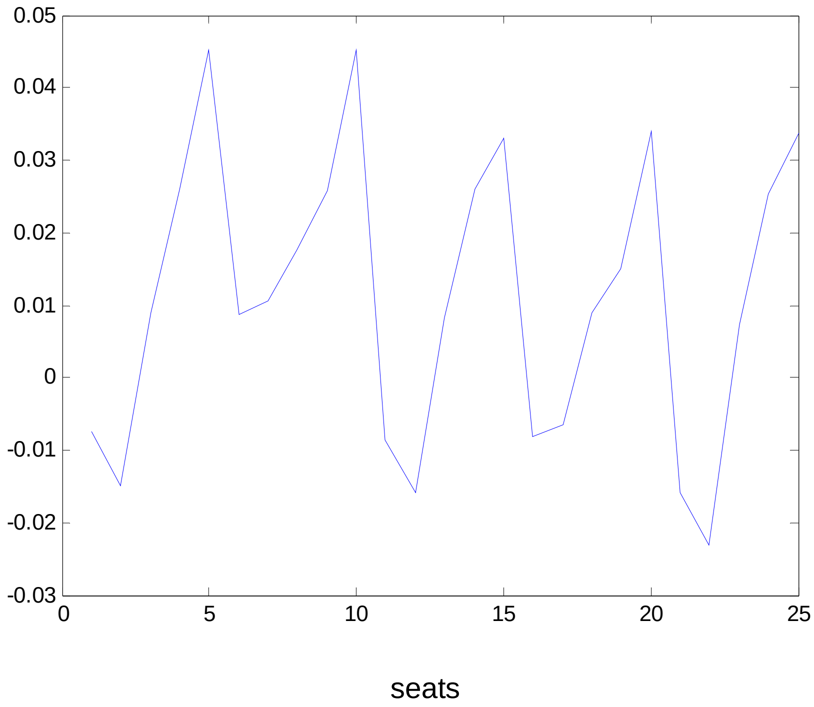


Figure 6.2 – Standard deviation between the trial data (input P_t and output T_t) and data in F = 0.1 (input P_t and output T_F).

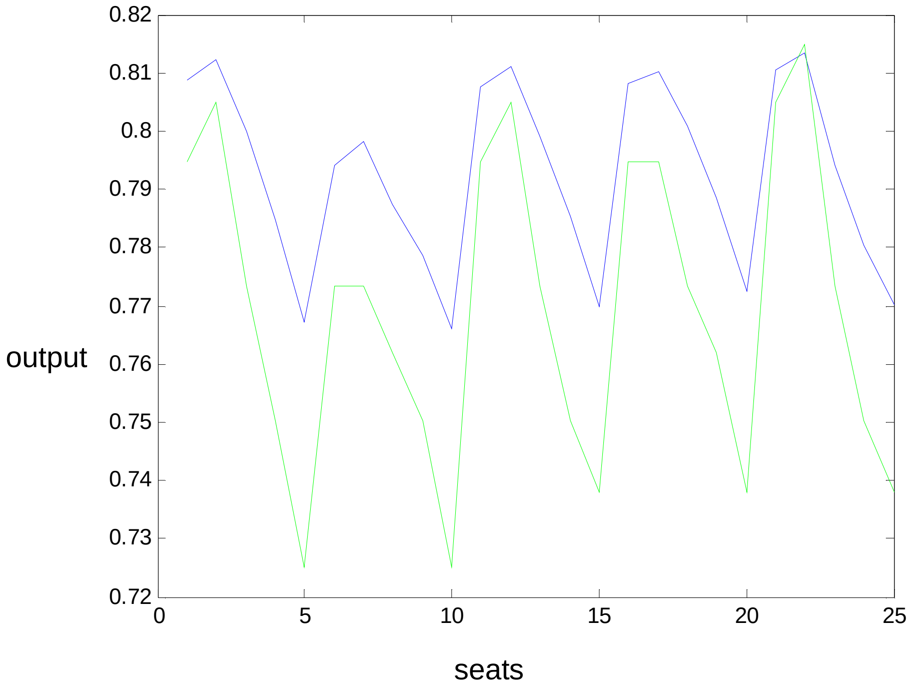


Figure 6.3 – Comparison of simulation results in F = 0.05. (Green – trial data, Blue – predicted data)

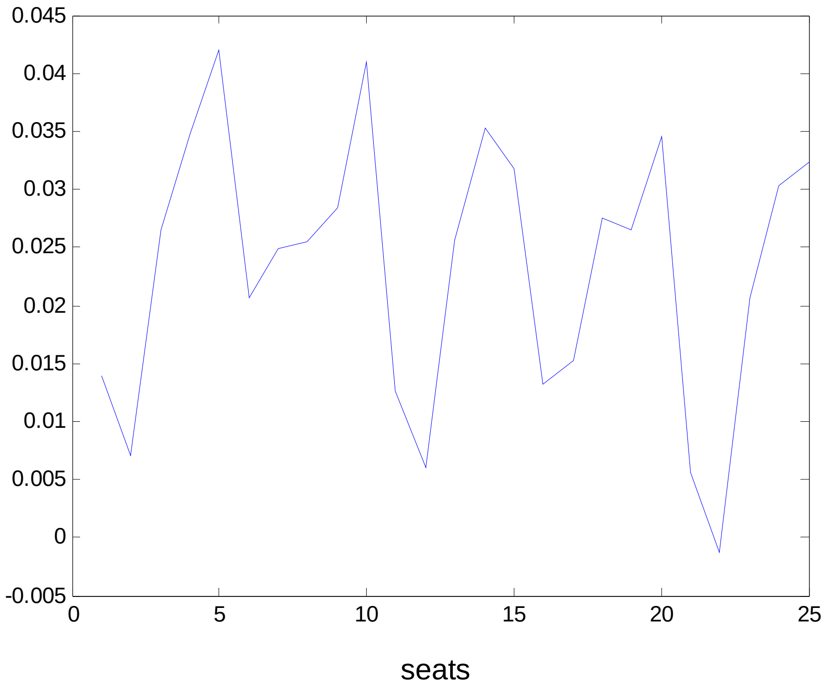


Figure 6.4 – Standard deviation between the trial data (input P_t and output T_t) and data in F = 0.05 (input P_t and output T_F).

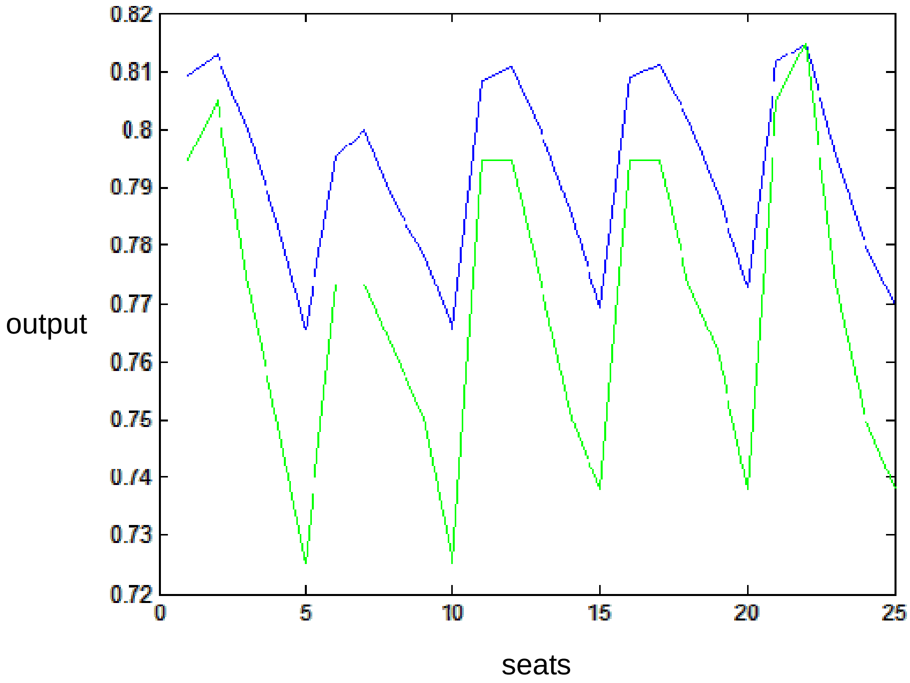


Figure 6.5 – Comparison of simulation results in F = 0.05 during verification process. (Green – trial data, Blue – predicted data)

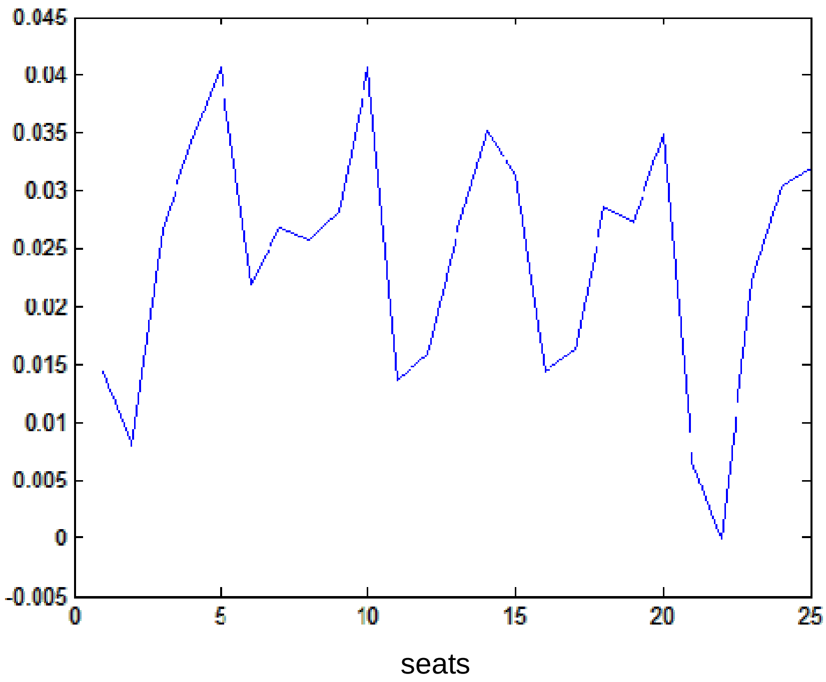


Figure 6.6 – Standard deviation between the trial data (input P_t and output T_t) and data in F = 0.05 (input P_t and output T_F) during verification process.

Chapter 7
Concluding Remarks

7.1 Conclusions

The objective of this study is to identify the factors affecting speech transmission inside a classroom in order to improve speech intelligibility. Through the understanding of the acoustic conditions inside a classroom, its acoustics characteristics and the most influential parameters are defined. However, quality of speech transmission is subjective to experiences and to human perception. The age, educational background, grew up environment, condition of health, knowledge and other experiences of the listener, also affect the human perception on understanding. The findings and results of this study are concluded in this chapter.

Material for experiments on this study has been firstly developed. Hong Kong has been selected as subject to carry out the experiments the study. Since Hong Kong is officially tri-lingual in education and in legal action, it is essential to have database of languages in Cantonese as speech for subjective speech intelligibility tests. The database of languages in Cantonese for speech intelligibility tests cannot be in Cantonese. There should be phonetically balanced monosyllabic words. In the first part of this study, 10 lists of phonetically balanced monosyllabic words in Chinese characters with Cantonese pronunciation have been created. There are 75 characters in each list. A statistical analysis has been conducted and compared with those in Mandarin and English. Since Cantonese is a tonal language, as opposed to English that is a non-tonal language; it helps to further the investigation on different languages effect on speech intelligibility. The most significant contribution of this chapter is to demonstrate a procedure that easily adapts for educators and researchers who are without knowledge of linguistics, to develop a database of other languages they use in their countries. With the help of the present study as a guideline for developing pertinent information in different languages, the speech intelligibility can be assessed easily by using those databases in other languages in subjective tests. Furthermore, when it comes to medical aspects, it is beneficial for assessing hearing impairment and developing hearing aid equipment.

After the groundwork of testing materials, a study on national and international standards and guidelines in classroom is carried out for comparison of the situation in Hong Kong. 20 classrooms are selected among primary schools, secondary schools and a university, for acoustical measurements. The selected 20 classrooms have similar geometrical configurations but differ in acoustics treatments. Acoustical parameters are studied and compared with those standards, which included indoor ambient noise level, outdoor noise level, noise criteria from mechanical ventilation systems inside the classrooms, reverberation time, speech transmission index and rapid speech transmission index. With references to the standards from thirteen countries and the recommendation from three organizations, indoor ambient noise level and reverberation time are crucial acoustical parameters to access the acoustical conditions of a classroom.

The results of field measurements showed that the acoustic conditions of most of the measured classrooms are unsatisfactory. Noise environment and large ambient noise level, long reverberation time and large mechanical equipment induced noise are comparatively low with requirements of standards and recommendations. In addition, the speech transmission index and its simpler form, the rapid speech transmission index, are low in values. These inadequate acoustics conditions lead to ineffective speech communication inside the classrooms. Especially the reverberation time, it is comparatively long in our measurements, far from an acceptable range. The unexpected results bring out a call to revise the standard design and to devise acceptable standards for classrooms in Hong Kong.

As the classrooms are different in acoustics features, the study further investigates the effect of absorption materials, such as acoustic back wall panel, acoustic ceiling and carpet. The speech transmission is apparently affected by absorptive materials, by their location of installation, area of the materials, absorption coefficients, absorption frequencies and other of their acoustical characteristics. The proprietary materials are designed to provide optimized sound absorption that will improve the acoustic environment of the classroom. Thus, selection of materials inside a classroom is important for improving speech intelligibility at design stage.

Application of acoustics simulation and modelling software has been studied for prediction on performance and impact inside a classroom under different types of acoustics treatment. From the result of simulation, it is recommended to install acoustic reflector and absorptive panels at the end walls and ceilings of classrooms, especially the acoustics ceiling. Their absorption characteristics to shorten the reverberation time are undoubtedly.

Subjective speech intelligibility tests are carried out to find the relationship between subjective hearing experiences and objective acoustical parameters. The signal-to-noise ratio, which is the most influential parameter in classroom acoustics, has a close correlation with PB word scores, which also is subjective perception related. The higher the signal-to-noise ratio, the better the PB word scores can be achieved by the subjects of different age groups. For different age groups investigated in this study, the young adults aged 19 to 22 years old achieved higher stability in PB word score under lower signal-to-noise ratio than the children. There are 2 possible reasons: 1) the hearing systems of young adults are well developed; 2) the knowledge and experiences grow and increase along with their age, it is easier for them to guess what they heard under complex listening environment. It is suggested that the signal-to-noise ratio should be higher than 11 dB(A) for over 70% of speech perception by PB word scores calculation. However, the stringent conditions in signal-to-noise ratio are vital for the earlier age of students.

It is also remarkable that the reverberation time is unacceptably long, even though acoustics ceiling and carpet have been used as proprietary material of acoustics treatment. The interior structure may be one of the reasons for high reverberation that a crossbeam cross over the middle at the ceiling leads to diffusion inside the classroom. Long reverberation time can interfere the continuous signal of the speech and therefore, masking signal and cocktail party effect like echo easily takes place, especially when surrounded by hard wall inside the classroom. This leads to lower the speech intelligibility. Nevertheless, the presence of acoustics ceiling and carpet are effective means to maintain an even distribution of signal-to-noise ratio, and helps t absorb the reverberant energy with their absorption coefficient and characteristics.

From the final chapter in this study, computerize and engineering know-how is applied for data modelling to predict and estimate the effect of each acoustical parameters to PB word scores. Through the application of a neural network, the experimental data from previous chapters is used for iteration on the learning process and modelling. Among three acoustical parameters: signal-to-noise ratio, reverberation time, and speech transmission index, signal-to-noise ratio is the most influential to PB word scores. The effect is crucial and it implies that signal-to-noise ratio is an indispensible factor in classroom acoustics. However, the effect of reverberation time to PB word scores is dramatically increased under the same value of signal-to-noise ratio. The impact of speech transmission index to the PB word scores is vague but their relationship clearly states that with higher STI, the PB word scores are higher. It is further understood that STI is an objective index to assess the quality of speech transmission, but not to affect parameter to PB word scores. It is suggested that signal-to-noise ratio and reverberation time are the important factors to affect the PB word scores, which is the speech intelligibility inside the classrooms.

As a conclusion, it is suggested that 1) the signal-to-noise ratio should be higher than 11 dB(A) for both tonal and non-tonal languages in order to achieve over 70% of speech transmission without the usage of address system; 2) re-design the interior structure, like removal of crossbeam at the ceiling to prevent any scattering effect; 3) windows and doors are recommended to use as sound insulators if the school is built in noisy environment, for example, surrounding by busy traffic or other noisy environment; 4) acoustics ceiling is recommended for acoustical treatment than carpets or acoustics back wall panel since its large surface area for absorption, and the most effective improvement on reverberant and other noise absorption by its absorption frequencies; and 5) selection of material inside the classroom should be aware to prevent high percentage of usage of hard wall. Furthermore, it is highly recommended to conduct the teaching in small classes, especially for young children, since the children in small classes teaching may receive more attention if the acoustics condition of the classroom is not acceptable. It is also avoided the effect of amplifying system to the hearing health of those young children.

7.2 Recommendations for future work

This study aimed to find out the relationship between objective acoustics parameters and subject hearing perception for improving speech intelligibility in the classroom as purpose. Nevertheless, certain aspects can be elaborated further, and extended in future research with similar approach.

Since there are many limitations in both administration and in practical, the classrooms selected and measured are only limited by very similar geometrical configurations. Besides, the number of classrooms is also very narrow; only 20 classrooms among 5 primary schools (though 2 classrooms from 6-th primary school for subjective tests), 5 secondary schools and a university are selected for the study. The number of classrooms and schools should be increased for further research to investigate further the effect of environment to the acoustical conditions inside the classroom, so to the effect of speech transmission.

Furthermore, the study in effectiveness of different acoustic treatment should be conducted in the classrooms. For more understanding on their performance with comparison on each treatment, it helps to shorten the reverberation time in order to improve the speech intelligibility with more practical fact as reference.

As the classrooms design in the universities is different, it is interesting to extend the study on the classrooms in the universities, with different geometrical characteristics, different interior design and different acoustics features inside the classrooms, ranging from small tutorial rooms for 10 students, to lecture theatres for about 300 students. Although this will not focusing on improving the speech intelligibility, it aims to investigate the effect of geometrical characteristics, different interior design and different acoustics features to acoustical parameters. It will then form a based as guidelines for design.

The current study on classroom acoustics or speech intelligibility by application of neural network is the case of our experiments. Although the results concluded a meaningful application and suggestion, however, the database is not sizable. The assistance of computerize and engineering know-how, use neural network as data modelling to predict and estimate the effect of speech transmission by objective acoustics indices: signal-to-noise ratio, reverberation time, early-to-late arriving ratio or useful-to-detrimental ratio, etc. is practical and applicable in real world. Development on computerize estimation on effect of those objective acoustics parameters to quality of speech transmission can minimize the time for thousand iteration on the computing. In addition, it can simplify the usage as a program embeds on the webpage for educators or school developers to understand and have earlier consideration on design stage.

To conclude, the current study and recommended future research should aim to benefit the society through improving the environment for education.