COMPARATIVE ANALYSIS OF ON-SITE STIPA MEASUREMENTS WITH EASE PREDICTED STI RESULTS FOR A SOUND SYSTEM IN A RAILWAY STATION CONCOURSE

Akil Lau[1] and Deon Rowe[1]

[1]Building Sciences,
Aurecon, Sydney NSW 2089, Australia
Email: akil.lau@aurecongroup.com
Email: deon.rowe@aurecongroup.com

Abstract

Houtgast and Steeneken designed the Speech Transmission Index (STI) test signals (modulated, speech-shaped noise) and its measurement methodology in 1970. The underlying basic principle of the STI is preservation of speech intelligibility during transmission. This is achieved by maintenance of the natural intensity fluctuations in the speech spectra. The design of the test methodology was such that they measured the natural modulations in a way that measurements could be carried out precisely and effectively, notwithstanding the constraints of the environment. Limitations associated with on-site STI testing accounts for certain deviations in the final results. Speech Transmission Index for Public Address (STIPA) is a simplified version of the standard STI method which is dedicated for sound level meters. STIPA measurements were conducted at a train station concourse by the Direct method (using speech signal modulation) and the Indirect method (using pink noise for impulse response) in accordance with [5]. A detailed acoustic model of the concourse was also created using EASE version 4.3 for STI predictions. The software Enhanced Acoustic Simulator for Engineers (EASE) uses MTFs (Modulation Transfer Functions) as a basis for high accuracy STI predictions. The software uses seven averaged frequency dependent octave band impulse responses to calculate the Modulation Transfer Index (MTI) values by means of 14 modulation frequencies. This paper evaluates the deviation in the measured STIPA results versus EASE-predicted STI results for the sound system in the railway station concourse and investigates the primary causes for the variation in the results.

1. Introduction

Transport for NSW (TfNSW) has engaged Novo Rail alliance to design and deliver the works for the Wynyard Station Upgrade project (WSUP), including speech intelligibility assessment. STIPA measurements and assessment was conducted at the concourse area of Wynyard Station in accordance with [5].

Three broad scenarios were assessed which are mentioned in Table 1. Comparison of STI for three scenarios helped to compare the speech intelligibility at the Wynyard concourse during the different stages of the upgrade and provided guidance on the internal surface treatments including suspended ceiling specifications for the upgrade. As the project is ongoing, we haven't documented the specifications of internal finishes, etc. in detail.

Table 1. Scenarios assessed

Scenario	Existing/ Future	Suspended ceiling
1	Existing	Yes
2	Existing	No
3	Future	Yes

The preservation of speech signal inside public spaces is essential for transmitting useful information to commuters and a requirement under Australian Standard for Fire Protection [2] with limits mentioned in Section 5. Speech intelligibility is a measure of how comprehensible speech is, or the degree to which speech can be understood. Intelligibility is affected by spoken clarity, explicitness, lucidity, comprehensibility, perspicuity and precision. Speech signal is usually degraded due to limitations imposed by the transmission channel. The degradation in the speech signal may result in modification of its frequency range, dynamic range, distortions, etc. On-site testing and predictions of speech quality is to quantify these limitations and to identify the physical aspects of the communication channel. The diagnostics and objective assessment methods can be categorised into three parts:

- Subjective measures based on a speaker and listeners;

Subjective tests make use of various types of speech signals. Frequently used speech items used for testing are phonemes, words (digits, alphabet, short words, etc.), sentences and free conversation in combination with quality rating.

- Predictive measures based on physical parameters;

Predictive measures are based on physical and perceptual parameters that quantify the effects on speech signal and loss of intelligibility due to a limited frequency transfer, masking noise, reverberation, echoes, non-linear transfer resulting from peak clipping, quantisation, or interruptions.

- Objective measures obtained by on site testing with specific test signals.

Specific measuring devices and test signals have been developed to objectively measure the speech intelligibility in a given environment.

An objective method to predict the speech transmission quality of an existing communication channel was developed by Houtgast and Steeneken [3] [4]. The transmission quality is derived from an analysis of the received test signal and is expressed by an index, the Speech Transmission Index (STI). The STI is based on weighted contribution from a number of frequency bands (125 Hz – 8000 Hz).

STI measurement requires a special test signal from which the effective signal-to-noise ratio in each octave band at the receiving side is determined. The design of the test signal allows an adequate interpretation of many degradations than just a limited frequency transfer and masking noise. Non-linear distortion and distortion in time domain are also accounted for with the special test signal.

2. Overview of STI Prediction Methodology

STI prediction is based on the contribution of a number of frequency bands (125 Hz – 8000 Hz which is specific to the frequencies of typical speech) determined by the effective signal-to-noise ratio. Signal-to-noise ratio is determined predominantly by the background noise level but also by the products of distortions in the time domain (reverberation, echoes, and automatic gain control) and non-linearity (e.g. peak clipping or quantization). The values of effective signal-to-noise ratio are limited to the range of -15 dB to +15 dB. The test signal consists of the frequency spectrum of speech signal. Each octave band is modulated with a periodic signal in such a way that the intensity envelope is modulated sinusoidal. The modulation only works in the intensity domain, with the distortions affecting the degree of modulations of a sine-wave without affecting the sine-wave shape, thus reducing intelligibility. Steeneken and Houtgast [4] suggested the relevant range of modulation frequencies extends from 0.63 Hz up to 12.5 Hz at 1/3-octave steps.

A long term speech spectrum noise is amplitude modulated by a signal which results in a sinusoidal intensity modulation. Octave band filtering and intensity envelope detection is applied to the signal. From the resulting envelope function a Fourier analysis determines the modulation index reduction, due to reduction in the transmission channel. This procedure is repeated for all the octave band frequencies and for all modulation frequencies. The extent of the reduction in the modulation of the original signal is quantified by Modulation Transfer Function (MTF).

The procedure mentioned above takes into account distortions such as band-pass limiting and noise masking as well as distortions in time domain. Non-linear distortions however act differently. If a speech signal passes through a channel with non-linear distortion, harmonic distortion and inter-modulation components will be generated in other frequency bands. Therefore modulation of each octave band with the modulation frequencies should be done separately. Otherwise, non-linear distortions components cannot be discriminated from the modulated test signal.

STIPA measurements consists of only one test signal with a predefined set to two modulations in each of the seven octave bands. The 14 modulations are generated simultaneously. There are two types of measurement systems that can be employed which are direct and indirect. Direct STI method uses modulated (speech shaped) test signals to directly measure the MTF. Indirect method uses the impulse response (or Pink noise) to derive the MTF. Direct STI method accounts for non-linear distortions, e.g. clipping, whereas the indirect method is only used for linear systems.

2.1 Direct method

STI can be measured directly using a suitable modulated signal played through the PA system to be tested. Direct method uses a test signal that has similar spectral and temporal properties to those found in a natural speech.

2.2 Indirect method

Indirect method of measuring STI corresponds to the mathematical manipulation (statistical method) of the system impulse response by means proposed by Schroeder applicable to linear and time-invariant systems. The impulse response give rise to MTF which subsequently calculates the STI. The following equation is used to calculate the MTF $m_{f,k}$ at modulation frequency f_m in octave band k:

$$m_k\,(f_m) = \frac{\left|\int_0^\infty h(t)^2 e^{-j2\pi f_m^t} d_t\right|}{\int_0^\infty h(t)^2 d_t} * \left[1 + 10^{-SNR/10}\right]^{-1} \quad (1)$$

$m_k(f_m)$ = Modulation transfer function of the transmission channel;
f_m = Modulation frequency
$h(t)$ = impulse response of the transmission channel;
t = integration variable for time;
SNR = signal-to-noise ratio in dB

Assuming diffuse reverberation field inside the space, impulse response can be written as:

$$h(t) = \frac{Q}{r^2} * \delta(t) + \frac{13.8Q}{r_c^2 T} * e^{-\frac{13.8t}{T}} \quad (2)$$

Q = Directivity factor for the sound source;
$\delta(t)$ = Dirac (delta) function;
r = talker to listener distance;
r_c = critical distance in the room or space;
T = reverberation time of the room or space.

Indirect method is closely related to the reverberation time (2) of the space. This equation does not show the relation of background noise in comparison with the reverberation time. Based on the previous studies [7] influence of background noise has an effect on the STI results for highly reverberant spaces (discussed ahead in Section 9.5).

MTF calculation used in statistical method (indirect) shown in (1) takes into account impulse response for the diffused space but fails to address the correction factor for auditory masking and hearing threshold (discussed ahead in Section 9.1 and 9.2).

3. Measurement methodology

STI measurement requirements are based on [5] summarised below in Table 2.

Table 2. STI measurement methodology summary

Attributes	Description
STIPA	Consists of a test signal with a predefined set of two modulations per octave band that are generated simultaneously giving a total of 14 modulation indices.
Distortions accounted for	Temporal distortion e.g. reverberation and echoes, Noise, Strong spectral distortion e.g. band-pass filtering.
Two methods to measure STI	Method 1 (Direct): Using modulated test signal (speech shaped sine wave)
	Method 2 (Indirect) : based on system's impulse response using Schroeder equation or pink noise (only applicable to linear, time-invariant systems)
Excitation signals allowed for STI measurements (Indirect)	TDS – Time Delay Spectometry (Only when background noise is low) MLS – Maximum Length Sequences (Only when background is low) Pink Noise Logarithmic sine sweep or Exponential sweep Speech shaped MLS without averaging – to measure the effect of background noise
Signal to noise ratio (SNR)	At least 20 dB (or 15 dB) should be obtained in all 7 octave bands (125 Hz – 8000 Hz).
Wind speeds during measurement	<4m/s Measurements using MLS are more vulnerable in this respect than measurements performed with sine sweeps.
Test signal	PCM/ WAV MP3 should be avoided though compression schemes employing at least 128 kbit/s have been shown to work without apparent error.
Test speech level	For measurements with a talker or other acoustic source, in the absence of a PA system, the test speech level shall be set to 60 dBA measured at 1m distance, on the axis of main radiation of the artificial mouth or talker. If it is required to simulate a condition with a raised vocal effort (Lombard effect), the test speech level shall be set to 70 dBA.
Number of measurements	If the influence of fluctuating noise cannot be reduced, measurements should be repeated at least 3 times before average STI.

4. Limitations with objective measures

The following limitations were a part of the objective measurements conducted at Wynyard railway concourse:

- Distortions in the measurement signal was avoided as the indirect method does not correctly account for the effects of distortion. Although existing ceiling speakers were not tested for distortions (beyond the scope of work), they were not used during testing with the indirect method.
- As the test signal is band-limited random or pseudo-random noise, repetition of measurements does not normally produce identical results, even under conditions of steady interference. The results centre on a mean with a certain deviation. The deviation depends, amongst other factors,

on the number of discrete measurements of the modulation transfer function (usually 98 for the STI method or 14 for STIPA) and the measuring time involved.

- With STIPA and a measurement time of 15 s, the maximum deviation is approximately 0.03 STI for repeated measurements. With fluctuating noise (for example, a babble of voices), higher deviations may be found, possibly with a systematic error (bias). This can be checked by carrying out a measurement in the absence of the test signal, which should result in a residual STI value less than 0.20. An estimate of the deviation should be made by repeating measurements for at least a restricted set of conditions.
- A worthy practice to average the STI results over two or three measurements for each location was followed for all the measurements conducted at the concourse.
- Due to limitation of safety and hearing impairment to commuters at railway concourse, STI testing using indirect method was conducted using an active loudspeaker by generating pink noise at SPL of above 100 dBA at 1m distance rather than generating pink noise through ceiling speakers.

5. Qualification bands

As per [5] nominal qualification bands for STI are given below.

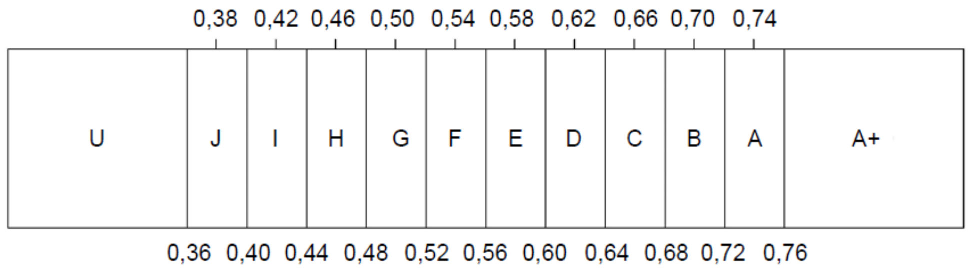

Figure 1. STI qualification bands with upper row of numbers are the STI values at the centre of the bands and lower row of numbers are STI values at the edges of the band.

6. Criteria

Based on the requirements in [2], the speech intelligibility of the loudspeakers for Emergency Warning & Intercommunication Systems (EWIS) within the evacuation zone should comply with the STI criteria mentioned in Table 3. The speaker system used at the railway concourse was EWIS over PA which basically indicates that same ceiling speakers were used for PA and EWIS system.

Table 3. Speech intelligibility criteria

Scenario	Speech Transmission Index (STI)
Normal power operation	≥0.5
Stand by power source operation	≤0.45

Based on [6], a minimum standard of 0.5 STI at all locations is required for any PA system. The target of STI should aim for 0.6 STI where practical, but is not mandatory.

7. Modelling methodology

Modelling of the STI in the concourse area with various ceiling materials was undertaken to compare the predicted STI ratings with the stipulated criteria using Enhanced Acoustic Simulator for Engineers (EASE) software. The software uses seven averaged frequency dependent octave band impulse responses to calculate the Modulation Transfer Index (MTI) values by means of 14 modulation frequencies.

8. Results

To legitimately compare the results from the objective measurements and predictive tools, Wynyard Station concourse without the suspended ceiling was modelled and tested (using direct and indirect methodology). Additional measurements with the upgrades at the concourse will be conducted in the future. The results of the comparison are summarized in Table 4 for STI and reverberation time in Table 5. The reverberation time measurements complied with [1].

Table 4. Comparison of STI at Wynyard concourse

Location	Objective measurement using Direct method	Objective measurement using Indirect method	Predictive analysis
Scenario 1 – Existing scenario with suspended ceiling			
A	0.53	0.54	0.57
B	0.49	0.54	0.59
C	0.39	0.56	0.58
D	0.40	0.61	0.59
E	0.32	0.60	0.57
F	0.36	-	0.56
Scenario 2 – Existing scenario without suspended ceiling			
A	0.44	0.36	0.41
B	0.42	0.40	0.44
C	0.31	0.31	0.43
D	0.29	0.49	0.44
E	0.31	0.61	0.42
F	0.28	-	0.44
Scenario 2 – Future scenario with suspended ceiling			
A	-	-	0.63
B	-	-	0.62
C	-	-	0.64
D	-	-	0.62
E	-	-	0.61
F	-	-	0.60

Table 5. Comparison of reverberation time at Wynyard concourse

	Average reverberation time (seconds)		
Location	Low frequency 125 Hz – 250 Hz	Mid Frequency 500 Hz – 1000 Hz	High frequency 2000 Hz – 4000 Hz
Scenario 1 – Existing scenario with suspended ceiling			
A	1.4	2.0	2.3
B	1.2	2.0	2.2
C	1.6	2.3	2.6
D	1.7	2.3	2.4
E	1.0	2.2	2.4
Scenario 2 – Existing scenario without suspended ceiling			
A	2.3	2.6	2.7
B	2.5	1.9	2.4
C	2.4	2.9	1.7
D	2.5	3.0	2.5
E	1.0	2.2	2.4

Table 6. Background noise measurements at Wynyard concourse

	Background noise measurements (dBA)				
Location	L_{Aeq}	L_{Amin}	L_{Amax}	L_{A10},	L_{A90},
A	62	59	77	63	60
B	58	56	66	60	57
C	63	53	82	65	55
D	66	59	99	69	61
E	74	56	103	76	65

Table 6 shows the background noise levels before the testing at each of the location. All the testing was conducted in the night time and early morning. Figure 2, Figure 3 and Figure 4 show the modelling results for the concourse for 3 scenarios discussed in this paper (refer to Table 1). Figure 5 shows the 6 locations at the concourse where on-site measurements were undertaken.

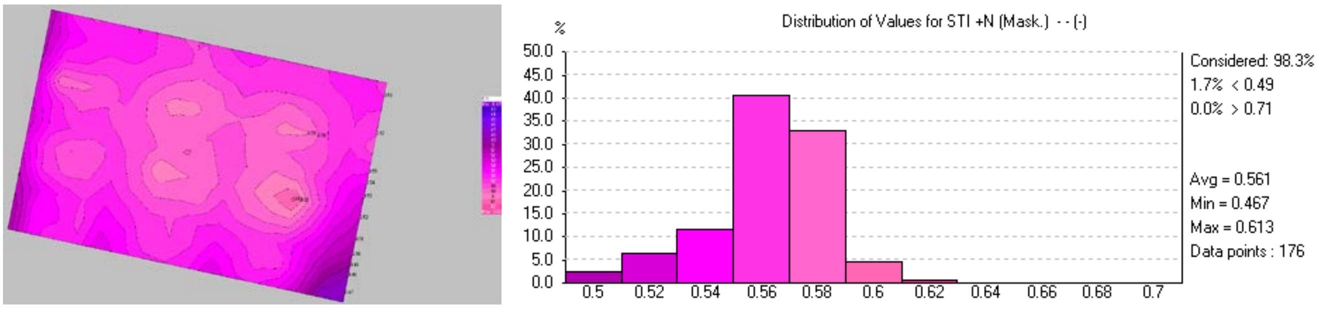

Figure 2. EASE model showing predicted STI at Wynyard concourse (Scenario 1)

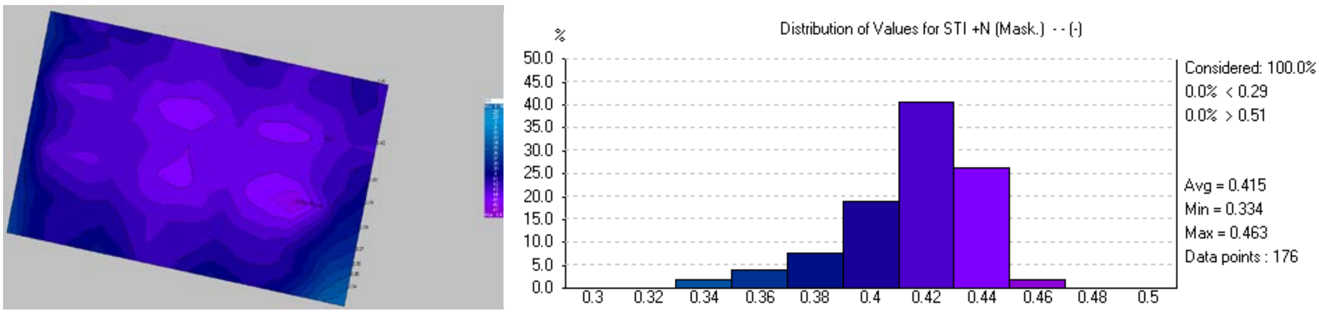

Figure 3. EASE model showing predicted STI at Wynyard concourse (Scenario 2)

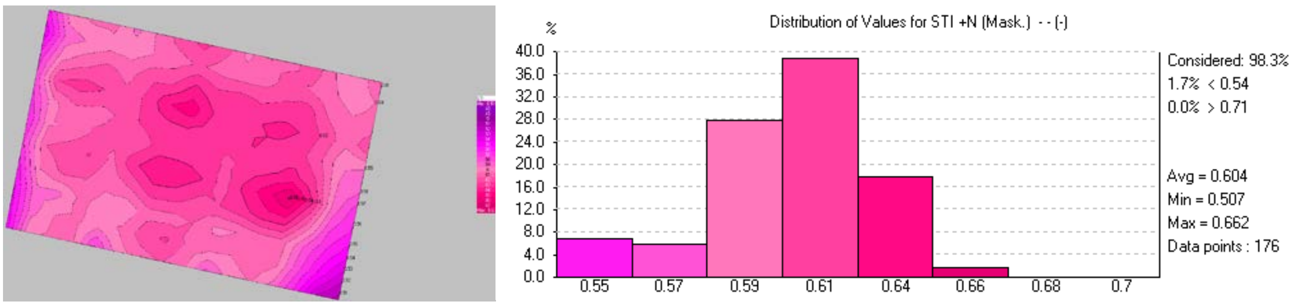

Figure 4. EASE model showing predicted STI at Wynyard concourse (Scenario 3)

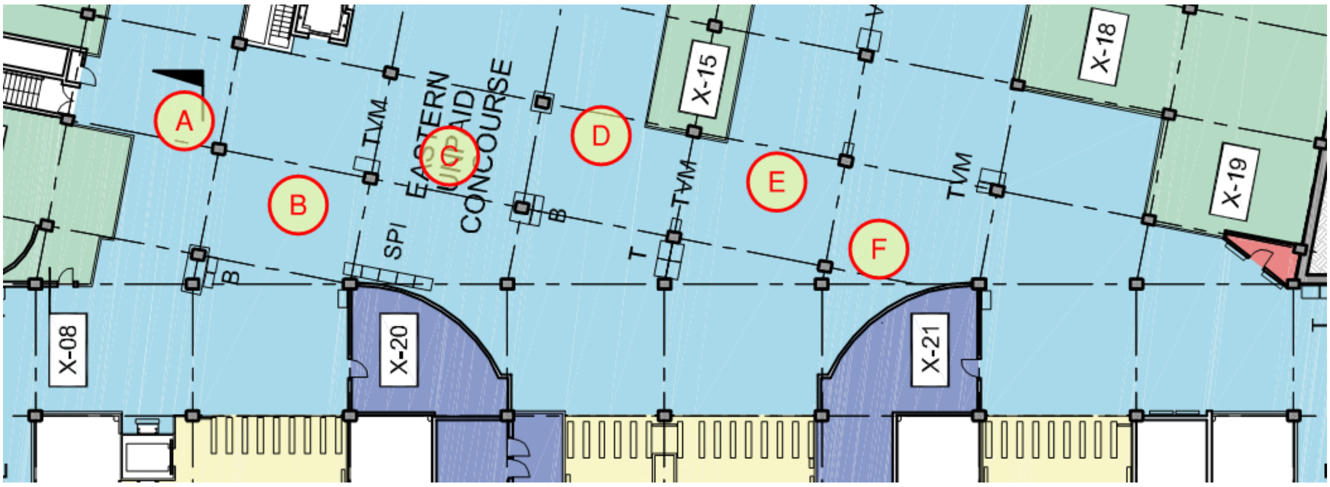

Figure 5. Plan of the concourse showing 6 location of measurements undertaken

9. Assessment

Below are the conceivable causes for the deviations in the results for the three methodologies employed to test/ predict STI at the station concourse.

9.1 Auditory masking

Besides the masking introduced by the noise in the transmission channel, an additional auditory masking phenomenon is inserted in the equation. Auditory masking is the effect introduced by human hearing organ where a strong masker in a lower frequency range reduces the perception of the tone or narrow band signal. The amount of masking depends on the level difference between the masker and the masked signal, on the absolute level of the masker and on their frequency distance. The masking effects are modelled as imaginary masking that decreases the effective signal-to-noise ratio, thus resulting in reduction of Modulation Transfer Index (MTI).
The masking intensity ($I_{am,f}$) for octave band K is given by:

$$I_{am,f} = I_{K-1} * amf \tag{3}$$

Auditory masking correction is introduced in objective measurement using direct method and also seen in predictive method but absent in the indirect method (calculation methodology in Section 2.2 shows absence of masking intensity). In absence of auditory masking taken into consideration, indirect method may give rise to higher STI values than expected.

9.2 Absolute hearing threshold

The absolute speech reception threshold is defined by the absolute threshold of hearing and the minimal required dynamic range for the correct recognition of speech. The effective signal-to-noise ratio is reduced when the speech level are low. The reception threshold intensity ($I_{rt,k}$) for octave band k is given by:

$$I_{rt,k} = 10^{\left(\frac{ART_k}{10}\right)} \tag{4}$$

ART_k = absolute speech reception threshold for octave band k in dB.
Similar to auditory masking, absolute hearing threshold adjustments are made in direct method and also in predictive method while absent in indirect method. Absence of the adjustments for hearing threshold will increase the effective SNR thus subsequently overestimating the STI values in the indirect method.

9.3 Reverberation time influence on indirect method

As seen in (2), reverberation time has a direct influence on the impulse response. Reverberation time measurements were undertaken inside the concourse at 5 locations using interrupted continuous noise source (Pink noise) as per [1]. Due to restriction of PA system for this testing, omni source loudspeaker was used to produce pink noise. Reverberation time is directly proportional to the amount of absorption on the internal surface of the concourse. As seen from Figure 5, there are negligible differences between all the 5 locations in terms of internal surface treatments which translated into the reverberation time results (with negligible difference as seen in Table 6) measured for the 5 locations.

Negligible difference in reverberation time gave rise to STI results with similar differences for all the 5 locations measured. A comparison of STI results with direct method shows major differences in location 4 and 5 (0.2 and 0.3 STI respectively) which were primarily due to poor performing PA speakers around that area.

Reverberation time for lower frequencies (125 Hz – 250 Hz) shows major variations in relation to with and without suspended ceiling scenarios (refer to Table 6 for scenario 1 & 2). The difference is evident for all the 4 locations and absent for location E. This was due to the fact that suspended ceiling near Location E was refurbished and replaced with metal tiles on grid (NRC ~ 0.01). Influence of reverberation time for lower frequencies (scenario 1 & 2) translated into corresponding STI values with similar differences (ranging from 0.25 – 0.12) between the two scenarios.

9.4 Fluctuating background noise

Indirect method (in comparison to direct method) of measurement are best suited for a situation where fluctuating noise is a feature (such as machinery repeatedly turned on/off, etc.). Wynyard concourse being a work site was subject to fluctuating noise during a few of the measurements (direct and indirect). Indirect measurement results are more reliable compared to the direct method in light of this information. This point contradicts the validity of direct method STI results over indirect method as per previous section 4.4.

9.5 Effect of source directivity on STI in high background noise environment

As per [7] source directivity has a substantial influence on STI in a highly reverberant soundfield as well as the existing background noise. Generally, sensitivity of STI to directivity reaches a maximum at relatively short source-receiver distance (less than 2m) for low background noise conditions and this sensitivity increases with reverberation time. As would be expected, a long reverberation time and long source-receiver distance yields low STI values. Introducing background noise makes STI less dependent on reverberation time, especially for larger volume rooms.

For large spaces background noise increases the sensitivity of STI to directivity for low source-receiver distance (approximately less than 2m), assuming reverberation time low. While the introduction of background noise decreases the sensitivity of STI to directivity inside large spaces for longer source-receiver distances (approximately greater than 5m). Both the above case assume the reverberation time is high and source sound power level is constant.

9.6 Temporal and spectral aspects of testing method

STI testing and prediction methods uses test signals which varies in temporal and spectral aspects compared with actual human speech. Below are some of differences:

- Absence of gaps in the test signal;
- Spectral differences between individual words and STI signal;
- Energy distribution in each time frame;
- Dynamic range of speech;
- Spectral differences between various talkers

9.7 Binaural hearing

STI measurements are based on monoaural listening as per [5]. Binaural listening for the purpose of STI estimation can have greater impact on the test results. Indirect STI measurements for the concourse were conducted using a single noise source while the ongoing construction work (cause of intermittent fluctuating noise) was located on a different direction to the source noise. Present measurement techniques did not take into account the direction of the predominant background noise source and test signal in lateral plane and its effect on the STI results. Humans have binaural hearing and currently there are no evident data on the influence of different location of noise source in the lateral plane compared to test signal location on STI results.

10. Conclusions

STI objective measurements and predictions were conducted for Wynyard railway concourse in order to compare the results and assess the cause for the deviations in the measured and predicted STI results. STI measurements were conducted in accordance with [5]. There are various causes that are highlighted in this paper which focusses on the causes for deviation in the STI results for large public spaces. The causes being auditory masking, absolute threshold shift, binaural hearing, reverberation time, fluctuating background noise, etc. Further investigations and research on test signals and STI measurement technique have to be undertaken to gain greater understanding on suitable testing methodologies for corresponding spaces.

References

[1] AS/NZS 2460:2002 Acoustics – Measurements of reverberation time in rooms, *Standards Australia*, Australia 2002.

[2] AS 1670.4-2004 Fire detection, warning, control and intercom systems – System design, installation and commissioning – Part 4: Sound systems and intercom systems for emergency purposes, *Standards Australia,* Australia 2004.

[3] Houtgast, T., and Steeneken, H.J.M. “Evaluation of speech transmission channels by using artificial signals”, *Acustica* **25**, pp.355-367, (1971).

[4] 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).

[5] IEC 60268-16:2011 Sound System equipment – Part 16: Objective rating of speech intelligibility by speech transmission index, 4th Edition, *International Electrotechnical Commission,* Geneva, Switzerland, 2011.

[6] RailCorp Digital Public Address System - Design and Installation Guideline, *RailCorp Guideline*, Version 1.0, 2013.

[7] Steward, K. and Cabrera, D. Effect of acoustic environment on the sensitivity of speech transmission index to source directivity, *Proceedings of Acoustics 2008*, Geelong, Victoria, Australia, 24-26 November 2008.