Audio Engineering Society
Convention e-Brief 310

Disagreement between STI and STIPA measurements due to high level, discrete reflections

Ross Hammond[1], Peter Mapp[2], and Adam J. Hill[1]

[1] *Department of Electronics, Computing and Mathematics, University of Derby*

[2] *Peter Mapp Associates*

Correspondence should be addressed to Ross Hammond (rosshammond@mail.com)

ABSTRACT

Objective measures of intelligibility, speech transmission index (STI) and speech transmission index for public address systems (STIPA), often form the basis for sound system verification. The reported work challenges the accuracy of both measures when encountering high level, discrete reflections. Tests were carried out in an anechoic environment with artificial reflections added between 0 and 500ms. Discrepancies were found to occur above 80ms due to synchronisation between modulation frequencies and reflection arrival times. Differences between STI and STIPA of up to 0.1 were found to occur for the same delay condition. Results suggest STIPA should be avoided in acoustic environments where high level, discrete reflections occur after 80ms and STI should only be used alongside other verification methods.

1 Introduction

Quantifying the intelligibility of public address systems (especially those used for voice alarm) is essential to verify that systems are capable of reproducing emergency messages that can be readily understood. An accurate method of measurement is significantly important, as an overestimation could seriously impact the safety of the public. The speech transmission index (STI) and its efficient, but slightly less accurate relative, the speech transmission index for public address systems (STIPA) are the most widely used objective measures of intelligibility for voice alarm and sound system design. Their entire methodologies can be found in the British [1] and International Standards [2].

This work challenges the accuracy of STI and STIPA, when used in environments or situations with high level, discrete reflections and intends to distinguish the difference between the two methodologies. These issues have been previously discussed [3] but this work intends to explore the specific differences between STI and STIPA to determine the potential errors and subsequently, the extent of the problem. Although it has been suggested that psychoacoustic factors mean the subjective impression differs from what STI scores reveal [3], the fundamental mechanistic issues are the focus of this work. Preliminary developments into a potential method of overcoming these issues are also briefly explored.

2 Background

STI operates by creating a modulation transfer function (MTF) matrix, formed of 98 modulation indices (MI), as seen in Table 1. MI values are obtained for 14 modulation frequencies across 7 octave bands. The individual MTF values can be captured individually (full STI) via measuring the modulation reduction of sinusoidally varying intensities or mathematically derived from an

impulse response (indirect STI) [4]. Reductions in MI occur due to ambient noise, reverberation or reflections. Each array of MI values is averaged for each octave band, forming its transmission index, which are all are applied with an intelligibility weighting function to produce the final STI value between 0 and 1, which aims to predict the intelligibility. STIPA allows the measurement to be obtained much faster (and is more-often used) by only including 2 MI values per octave band, distributed as seen in Table 1.

		Octave Band (Hz)						
		125	250	500	1k	2k	4k	8k
Modulation Frequency (Hz)	0.63			✓				
	0.8						✓	
	1		✓					
	1.25					✓		
	1.6	✓						
	2					✓		
	2.5							✓
	3.15			✓				
	4						✓	
	5		✓					
	6.25					✓		
	8	✓						
	10				✓			
	12.5							✓

Table 1. MTF Matrix with ticks representing values used in STIPA measurements.

Previous work [3] demonstrates the mechanistic issue with using STI measurements in environments with high level, discrete reflections beyond ~80ms. This is due to the synchronisation between reflection 'delay' time and the distribution of modulation frequencies. When the modulation response is viewed in higher resolution, it is clear that 14 'samples' do not accurately represent the response for higher delay times (Fig 1).
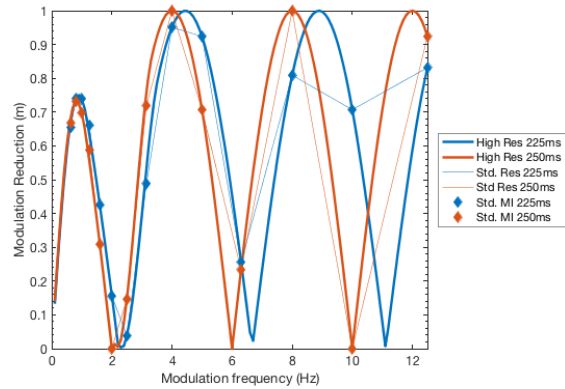

Figure 1. High resolution MTF compared with 14 MI values for 225ms and 250ms [3].

STIPA could possess further errors due to it only including two modulation indices for each frequency band. Averaging all transmission indices will be equal to that of STI, but since these values are weighted, their distribution will affect the result. Furthermore, when considering reflections will not be frequency independent, essential parts of the modulation response will be excluded from STIPA results. This concept can be seen in Fig 2, which shows the theoretical differences between modulation frequencies for each frequency band for a 250ms delay.
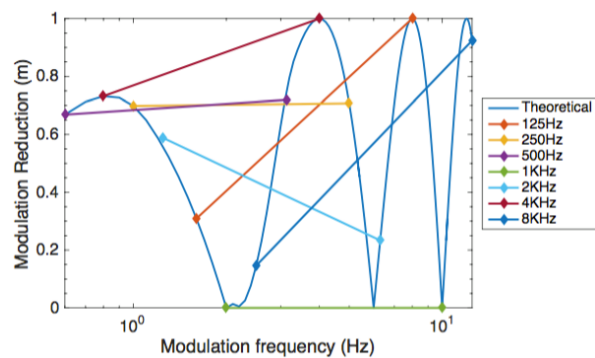

Figure 2. Modulation frequencies for STIPA with a 250ms delay [3].

3 Methods

To demonstrate the effects of discrete delays on STI and STIPA measurements, theoretically calculated data was obtained and compared with indirect measurements.

The theoretical STI, STIPA and MTF were mathematically calculated using Matlab [5], when the signal is summed with an exact duplicate at a given delay time between 0ms and 500ms. The interference between each modulation and delay time could be computed, allowing a modulation transfer function matrix to be constructed. For STI, this procedure excluded frequency dependence, since the modulation response would be identical for each frequency band. By averaging the corresponding two MI values for each STIPA frequency band, allowed the 7 transmission index values to be applied with the intelligibility weighting function required to generate STIPA data. By creating a MTF matrix, the theoretical STI and STIPA value, the result could be plotted for each delay time and compared.

Indirect STI, STIPA and MTF results were measured to be compared with the calculated data. Impulse responses were measured in an anechoic chamber at the University of Birmingham. Measurements were made with the Clio 10 software/hardware package by Audiomatica [6], via maximum length sequence (MLS). Extracted impulse responses were analysed with EASERA [7]. The MLS signal was distributed to two full-range, active loudspeakers placed equidistant to the omnidirectional measurement microphone at a distance of 2 meters. A Behringer X32 digital mixing console [8] allowed the signal to be sent to the two loudspeakers with delay applied to one. Both loudspeakers were set to a reference level of 65dBA.

Additionally, a Matlab script was created which theoretically calculates the STI and MTF for defined delay times, but incorporates additional modulation frequencies. The additional modulation frequencies were weighted according to the original spacing by creating intervals between the existing frequencies. This allows a 'high resolution' modulation response to be created and a high resolution STI result to be

calculated for each delay time. For the purpose of this investigation, the 14 modulation frequencies was increased to 131.

4 Results

The following figures (Fig 3 and 4) represent the MTF for two delay times, displaying both the mathematically calculated and measured data. The similarities provide evidence of a valid method of calculating MI values.
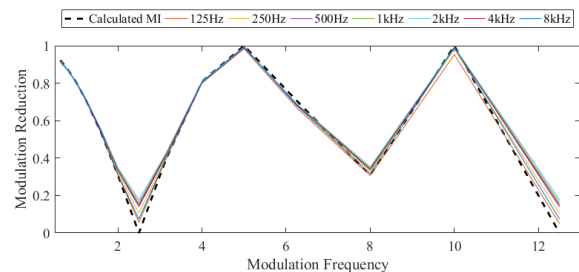

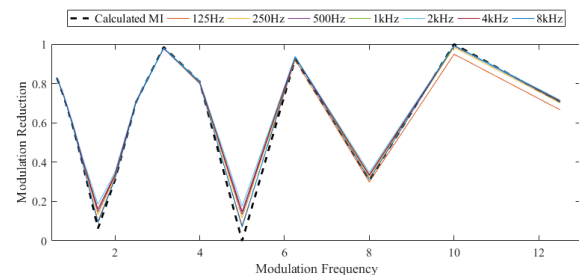Figure 3. Calculated and measured MTF values for a 200ms delay time.

Figure 4. Calculated and measured MTF values for a 300ms delay time.

The MTF values could be used to generate the theoretical STI and STIPA data for each delay time. This was compared with measured data for validation. Fig 5 displays calculated and measured STI and STIPA results which demonstrates how the reduced number and distribution of modulation frequencies for each frequency band for STIPA can

seriously alter the result, with a greatest difference of 0.112 and a potential overestimation of 0.0687. Fig 6 shows the theoretical difference between STI and STIPA for delay times between 0ms and 500ms.
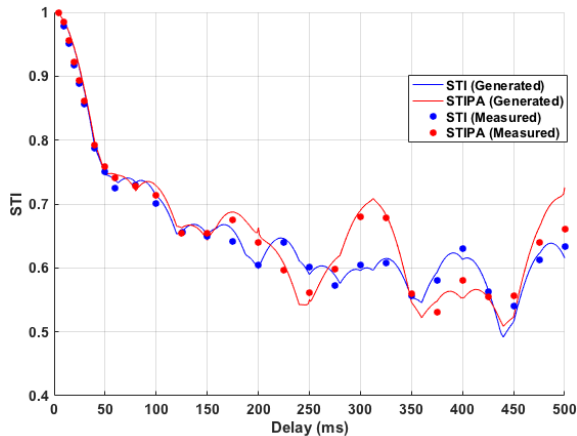

Figure 5. Calculated and measured STI and STIPA results for delay times between 0ms and 500ms.
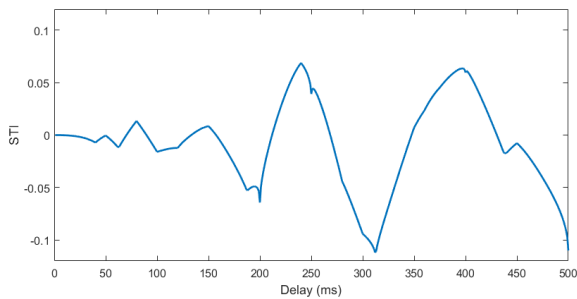

Figure 6. Differences between STI and STIPA for delay times between 0ms and 500ms.

Fig 7 demonstrates STI and STIPA compared with the modified STI score when the number of modulation frequencies is increased to produce a 'high resolution' response. A jagged curve begins to appear for longer delay times suggesting a greater number of modulation frequencies is required. However, the overall response is in line with expectation. The 'ripple' in the response is due to the two modulation frequency boundaries, where the

starting modulation is at 0.63Hz and the overall amount of deconstructive interference changes with delay. Starting below this point would be counterintuitive, as modulations below this do not contribute to intelligibility [1].
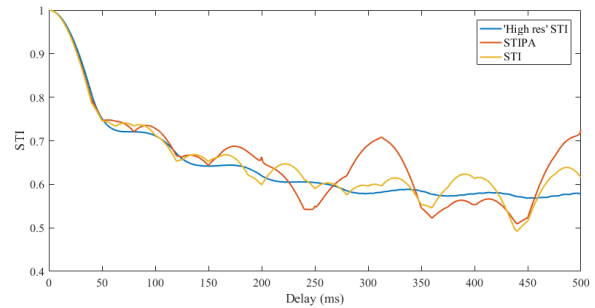

Figure 7. A comparison between STI, STIPA and 'high resolution' STI for delay times between 0ms and 500ms.

Comparisons with STI and STIPA allow the error in both to be quantified for each delay time. These errors can be found in Fig 8.
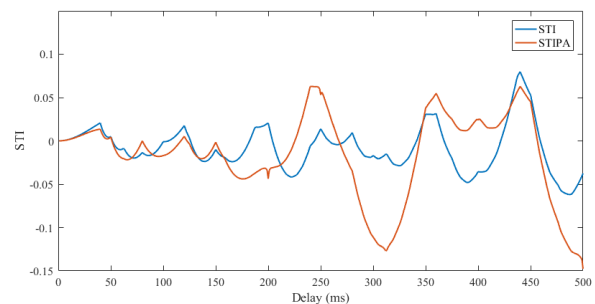

Figure 8. Total errors in STI and STIPA for delay times between 0ms and 500ms.

5 Summary and Recommendations

Building on previous work [3], results further demonstrate that STI inhibits mechanistic issues in environments with high level, discrete reflections beyond 80-100ms due to synchronisation between modulation frequency and reflection delay, whereas

STIPA is also affected by the distribution of weighted modulation frequencies.

STI can exhibit up to a 0.08 difference compared with a high resolution modulation response which better represents the acoustic environment, when delays are only considered up to 500ms. STIPA can exhibit up to a 0.112 difference compared with STI, and up to a 0.148 difference compared with a high resolution modulation response.

Therefore, it is recommended that STIPA should be avoided in this type of distortion and STI should not be the sole verification method, where temporal effects also need to be considered. Although designers may currently need to contractually oblige with STIPA measurements, findings indicate that at the very least, additional verification methods should be used.

6 Future Work

An indirect, high resolution MTF method would represent a room's acoustical characteristics with a higher degree of accuracy. However, this would completely alter the STI methodology which would require an extensive validation period. In reality, environments with this type of distortion do not appear regularly enough to condone modification. However, when the discrete delay time is known, the error produced in the STI score can be calculated. This could be implemented as a 'correction factor'.

Work has begun to determine the feasibility of this method. The error produced thus far has only incorporated a single, frequency independent delay with an identical level to the direct sound. In reality, discrete reflections will be frequency dependent, which will inhibit varying levels. Work has begun to determine a correction factor which will incorporate all frequency bands independently and also assess the direct sound level to discrete reflection level to noise floor ratio.

References

[1] British Standards Institute, "60268-16: Sound system equipment – Part 16: Objective rating of speech intelligibility by speech transmission index," 2011.

[2] International Electrotechnical Commission, "60268-16: Sound system equipment – Part 16: Objective rating of speech intelligibility by speech transmission index," 2011.

[3] Hammond, R., Mapp, P. and Hill, A.J., "The influence of discrete arriving reflections on perceived intelligibility and speech transmission index measurements" In: Audio Engineering Society Convention 141, 2016.

[4] Schroeder, M., "Modulation transfer functions: Definition and measurement," Acta Acustica united with Acustica, 49(3), pp. 179–182, 1981.

[5] Mathworks, "https://uk.mathworks.com," 2017.

[6] Audiomatica, "http://www.audiomatica.com," 2016.

[7] AFMG, "http://easera.afmg.eu," 2016.

[8] Music-Group, "www.music-group.com/p/P0ASF," 2016.