

ON THE IMPORTANCE OF THE SPEECH SPECTRUM FOR STI CALCULATIONS.

L Morales The Acoustics, Group, London South Bank University.
G Leembruggen ICE Design, Acoustic Directions, University of Sydney

1 INTRODUCTION

The ability of the Speech Transmission Index (STI) to predict the speech intelligibility of a transmission channel under noisy conditions is highly dependent on the assumed spectrum of the speech signal. The STI method (IEC 60268-16) specifies a standard male speech spectrum for STI calculations. Examination of the literature shows that the long-term average speech spectrum of male talkers differs substantially from the standard STI speech spectrum.

To explore these issues, the long-term average speech spectrum of forty male British English speaking people was first measured. Then, using several sound systems, the influence of this measured spectrum on the STI calculations was assessed, and compared with the standard speech spectrum.

2 BACKGROUND

The use of a specific speech spectrum has been recommended in ANSI standards since 1969 for calculations of objective measures of speech intelligibility. The speech spectrum recommended in ANSI S3.5 1969¹ (R1986) *American National Standard methods for the Calculation of the Articulation Index*, was mainly based on Dunn & White (1940) and French and Steinberg (1947) spectra. The spectrum recommended in a later version of the standard, ANSI S3.5 1997² (R2007) *American National Standard methods for the Calculation of the Speech Intelligibility Index*, was derived from several works which were compiled by Pavlovic³ in 1987.

The IEC 60268-16 standard introduced in 1998 gave gender-specific test signals for male and female talkers for use with STI calculations. These recommended spectra remained unchanged in later revisions of the standard, published in 2003 and 2011. However, no information has been provided in this standard about the source of the spectra.

Figure 1 shows speech spectra proposed by various sources. The spectra from French and Steinberg⁴, Benson and Hirsh⁵ (included in Pavlovic³ 1987), ANSI S3.5 1997² (R2007) and the speech spectrum for male speakers proposed by IEC 60268-16:2011⁶ are included in the figure. All spectra have been normalized to a level of 0 dB at 1 kHz. It can be seen that the IEC curve attempts to be a possible “best fit” to the other spectra for frequencies above 400 Hz. This “best fit” observed and the slope of -6 dB per octave suggest that a) the IEC curve tried to satisfy the speech spectra proposed in the ANSI standard (S3.5 1997), and b) the chosen spectrum was convenient since an slope of -6 dB per octave would be easy to obtain with an analogue filter.

Most of the published data for long-term averaged speech spectra (LTASS) are for the English language as spoken in the USA^{4,5,7,8}, Australia⁹, England¹⁰, and other countries¹¹. Although small differences have been found among English as spoken in different countries and among different languages, these differences are unclear as large difference among individuals have been found.^{12,13,14}

Cox and Moore⁸ (1988) concluded in their studies that the variability in spectra found among several languages was mostly due to individual talker differences. Additional research presented by Pavlovic et al.¹⁵ (1990) indicated that there was no difference in the speech spectra obtained for several languages.

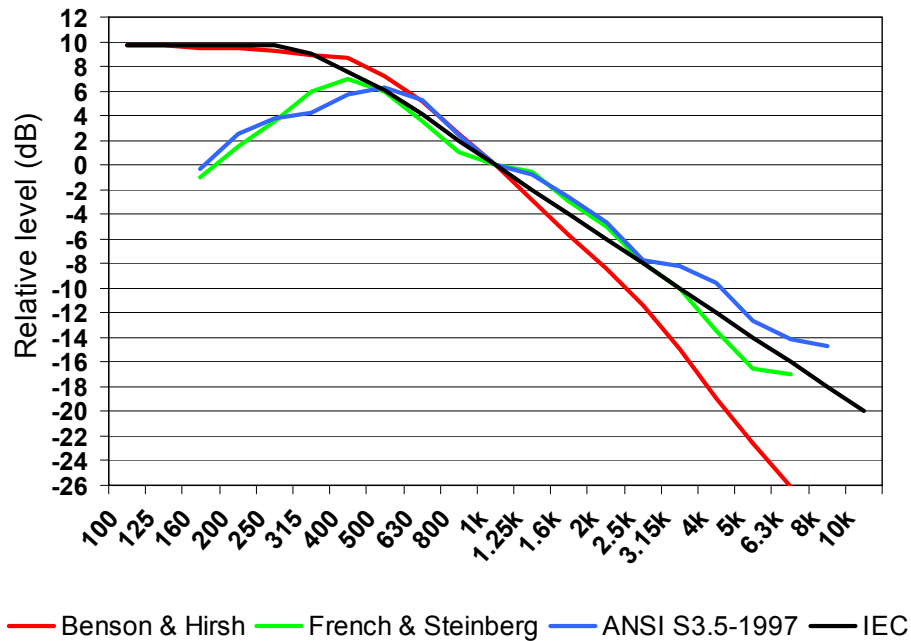


Figure 1. Speech spectra from different sources.
All the spectra levels are normalized to 0 dB at 1 kHz.

Although the LTASS is expected to be influenced by the type of speech material that is analysed, early studies of Benson and Hirsh⁵ (1953) concluded that the choice of speech material is not critical provided that is not grossly unrepresentative phonemically, such as speech passages containing repetition of a few phrases. It is also been established that the LTASS varies with vocal effort^{11,16,17} and that it is different for men and women.

The most extensive data on LTASS was provided by Byrne and colleagues¹⁸ in 1994. In their work, speech spectra for twelve languages were analysed. A passage from a story book was used for the recordings from which 64 s of speech were analysed. The microphone was placed in front of the talker at 20 cm in the same horizontal plane as the mouth and at an azimuth of 45° incidence, relative to the axis of the mouth. The azimuth angle of 45° was employed to avoid artefacts related to close talking conditions. Thirty-two talkers participated in the recordings for the British English language of which at least 10 were male. Although recordings for some of the languages were anechoic, no information was provided of whether the recordings for the English language took place in anechoic conditions.

Comparisons with the speech spectrum measured in front of the mouth (0°) have found a drop in high frequency content when the microphone was placed at 45° horizontal-incidence relative to the axis of the mouth^{19,20}.

3 EXPERIMENTAL DESIGN

Noting the effects on spectra associated with the different microphone techniques used in the literature, an experiment was designed to obtain the male British-English LTASS in anechoic conditions. The average long-term spectrum of conversational speech was recorded under anechoic conditions for forty male native British-English speakers with a microphone positioned in front of the subject’s mouth. Several microphone-to-mouth distances and different speech material were employed in the recordings. The measurement results were then compared with those in the literature.

A B&K 4188 microphone was used for the recordings and was connected to a B&K 2236 sound level meter (SLM) through an extension cable, with the SLM output connected to a Marantz PMD 661 recorder. The subjects’ speech was recorded into a wav format at a sampling frequency of 44.1 kHz and 16 bits resolution. The microphone was placed at the same height as the subjects’ mouths.

3.1 Subjects

Forty male native English subjects participated in the recording test. Figure 2 shows the age distribution of the subjects used. No subject reported any hearing or speech impairments. Ethical approval was sought and consent was given prior the experiments.

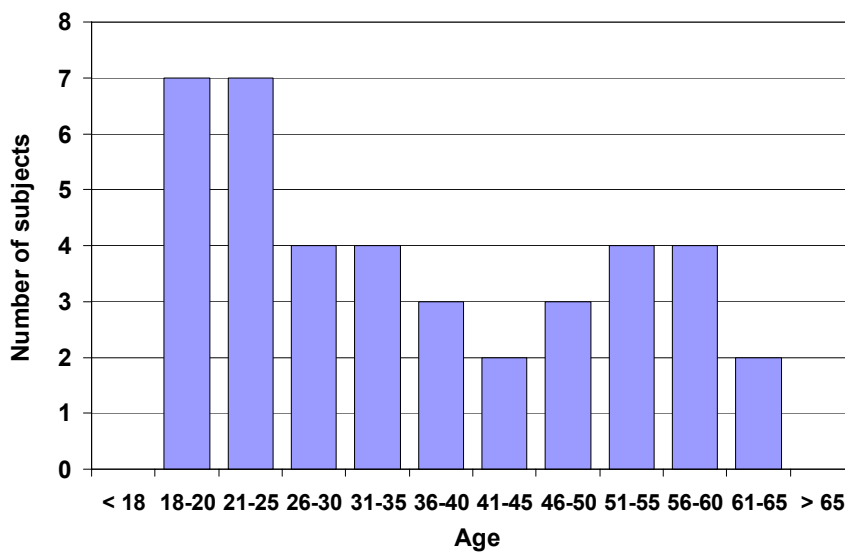


Figure 2. Age distribution of subjects.

3.2 Speech material

Three Harvard phonetically-balanced sentence lists extracted from the *IEEE Recommended Practice for Speech Quality Measurement*²¹ were used for the recordings. Each Harvard list contains 10 sentences incorporating specific phonemes at the same frequency they appear in English.

The three selected Harvard lists provided a few seconds more than the required 64s of speech¹⁹. The sentences were written on a thin sheet of cloth positioned in front of the subjects located at approximately 3 m from the microphone. The impact of the cloth on the measured spectra was investigated at the microphone location and found to have negligible effect.

The sentence “Joe took father’s shoe bench out, she was waiting at my lawn” was also used for the recordings. This particular sentence was chosen as it has been widely used in the previous literature and assumed to have spectrum similar to conversational speech^{3,5}. This assumption was found to be reasonably correct, although at some frequencies the agreement was less satisfactory³.

3.3 Procedure

The subjects were seated in a fixed chair placed within the anechoic chamber and given instructions to sit straight with their backs against the chair and to direct their voice towards the microphone. Prior to the recordings, subjects held a short conversation with the researcher to allow them to adjust to the anechoic environment. This adjustment period was intended to discourage the subjects’ from involuntarily raising their voice to compensate for the lack of reverberation and the extremely quiet conditions.

The participants were instructed to read the sentences from the cloth sheet at a normal speed and to maintain a constant speaking level. The subjects repeated sentences in which they made a mistake or in which they felt their fluency was unsatisfactory. The microphone was placed at 0.5 m from the subjects’ mouths for the Harvard sentence recordings, and at 0.5 m, 0.2 m and 1.0 m for the sentence “Joe took... waiting at my lawn”,

3.4 Spectral analysis

The spectrum of each talker was analysed using the waterfall function in the acoustic analyser WinMLS2004. The weighted overlapped segment averaging (WOSA) or Welch’s procedure²² was used with 1 s long Hanning windows and 50% overlap with a resolution of 1.0 Hz. For each waterfall slice, one-third octave band spectra were obtained from 100Hz to 10 kHz. All spectra were corrected for the response of the B&K 4188 microphone used for the recordings.

3.5 Results and discussion

The LAeq level of each talker was found by first removing the gaps between sentences using Audition v3.0 and then feeding the edited speech electrically into a B&K 2260 sound level meter.

The forty talker levels with the Harvard sentences were arithmetically averaged and found to be 63.8 dBA with a standard deviation of 2.7 dB, which was 0.6 dB higher than the levels obtained with the sentence “Joe took...at my lawn”. These measured levels are in agreement with current data reported for vocal efforts in anechoic conditions¹⁷.

For each subject, the measured 1/3rd octave band spectra were adjusted to normalise their overall level to 64 dBA. Figure 3 shows the speech spectra results for the 40 subjects with the Harvard sentences measured at the distance of 0.5 m.

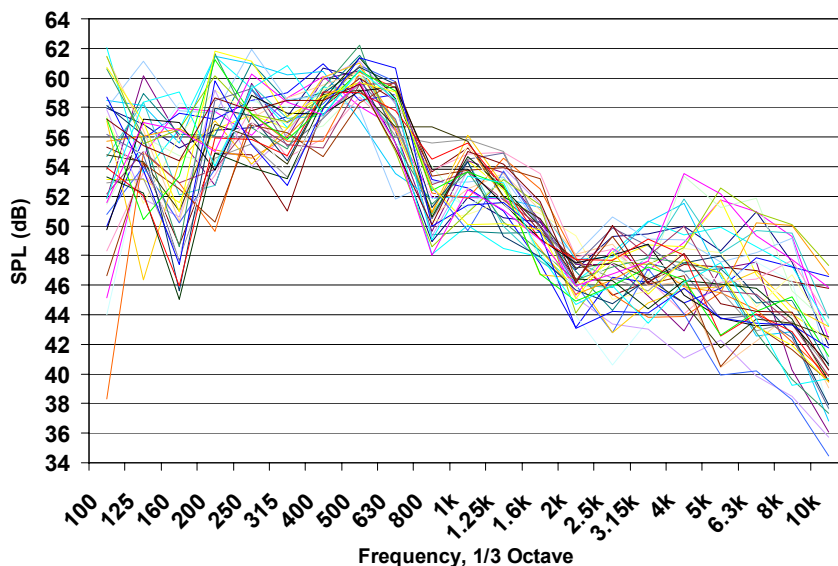


Figure 3. LTASS for the 40 subjects measured with Harvard sentences at 0.5 m. Each subject’s spectrum was normalised to 64 dBA.

The LTASS results presented in Figure 3 showed a high variability among subjects, which is in agreement with previous works found in the available literature.

Figure 4 shows the average spectrum obtained for the 40 subjects with Harvard sentences at 0.5 m. The IEC male spectrum⁶, the ANSI S3.5 1997⁴ (R2007) spectrum and the male spectrum from Byrne et al.¹⁸ were also plotted. Each of the 1/3rd octave bands was calculated on an average basis^{3,18}. All spectra values were normalised to an overall broadband level of 64 dBA. Tabulated values can be seen in Table 5 of the Appendix.

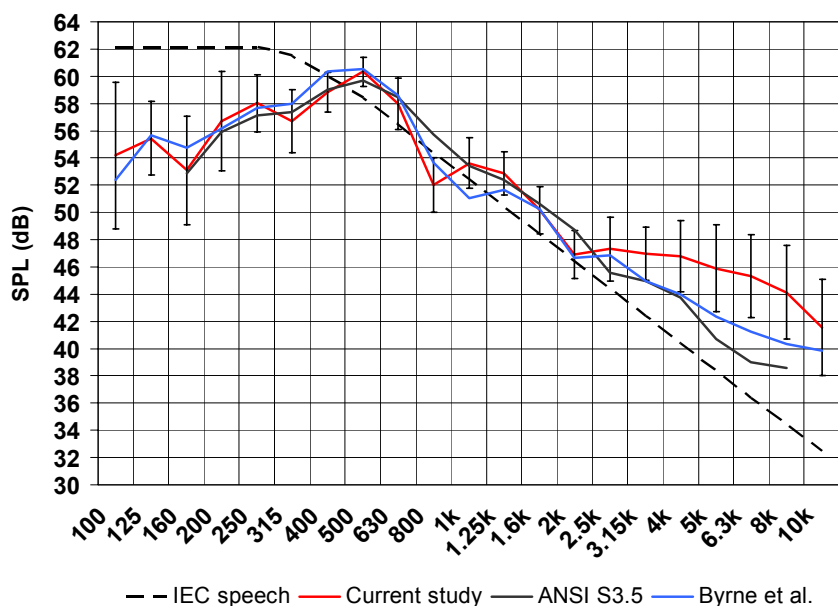


Figure 4. LTASS for this study measured at 0.5 m with Harvard sentences. Error bars indicate one standard deviation. Spectra from other sources were also included. All spectra were normalised to an overall level of 64 dBA.

The LTASS results presented in Figure 4 for the current study showed a good agreement with the IEC male spectrum for frequencies between 400 Hz and 2 kHz, and large differences were found for frequencies outside that range.

The values obtained in the current study were in agreement with the ANSI spectrum and the male spectrum presented by Byrne et al. for frequencies below 2 kHz. For frequencies exceeding 3 kHz, the current study presented between 2 and 4 dB more energy than Byrne et al., and between 2 and 6 dB more than the ANSI spectrum. The ANSI 1997 was derived from Pavlovic 1987, who averaged the spectra of six workers. However, the exact number of male talkers could not be determined from the references cited by Pavlovic 1987.

Byrne et al. mentioned an expected decrease of 2 dB for frequencies of 1 to 5 kHz and even greater decrease at frequencies above 5 kHz, due to the directivity of the talkers and the placement of the microphone at 45° incidence. Although previous research on the directivity of human speech in the horizontal plane showed a similar decrease at mid and high frequencies for 45° incidence¹⁹, other workers reported slightly lower directional losses²⁰. We conclude that the 0° incidence of the current study can mostly explain the differences with the Byrne et al. spectrum at mid and high frequencies. It is also noted that the average spectra of the UK male talkers presented in Byrne et al shows between 1 and 3 dB higher levels in the 2.5 kHz to 5 kHz region than the world average.

It is also unclear whether the recording method and/or a potentially lower number of subjects used for the ANSI spectrum were responsible for the differences in spectra with the current study.

Figure 5 (a) shows the difference between the LTASS obtained with Harvard sentences and the sentence “Joe took ... at my lawn”, while Figure 5 (b) shows the differences between the LTASS measured at 0.2 m, 0.5 m and 1.0 m using the sentence “Joe took ... at my lawn” .

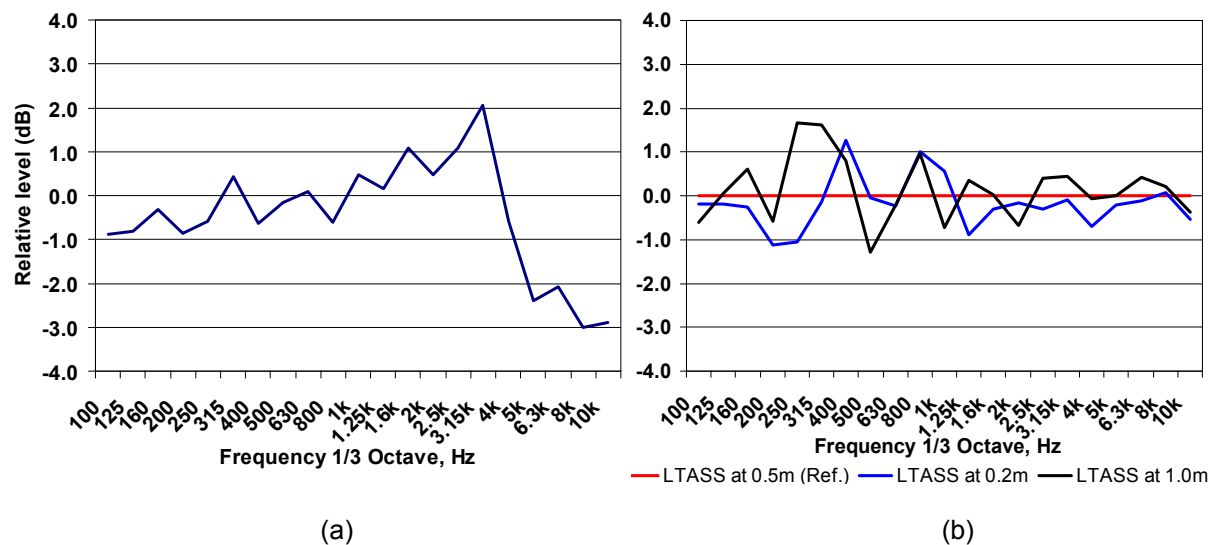


Figure 5. a) Difference between the LTASS obtained with Harvard sentences and with the sentence “Joe took...at my lawn”, both measured at 0.5 m. Positive values indicate higher levels with the sentence “Joe...”. b) Difference between LTASS measured at 0.2 m, 0.5 m and 1.0 m with the sentence “Joe took ... at my lawn”.

The results presented in Figure 5 (a) show good agreement at frequencies below 2 kHz between the LTASS of the Harvard sentences and the sentence “Joe took...at my lawn”. The agreement was less satisfactory for frequencies above 2 kHz. Figure 5 (b) shows that the distance at which the recordings were made did not influence the spectra substantially. The variation in the LTASS with recording distance is lower for frequencies above 1.6 kHz frequencies than for frequencies below 1.6 kHz. It is also noted that the variation in the LTASS with microphone distance is less than the measured variation in talkers.

The male spectrum obtained with Harvard sentences at the distance of 0.5 m is expected to better represent a LTASS for male British English speakers for use with sound systems.

3.6 Proposed male spectrum.

Table 1 shows the normalised values proposed in this study for the LTASS for British male English speakers, along with the male-talker equivalents given in IEC 60268-16. Both the LTASS and IEC spectra are normalized to an A-weighted level of 0 dB for easy scaling.

Octave band Hz	125	250	500	1k	2k	4k	8k	A-weighted
Male (IEC-60268-16)dB	2.9	2.9	-0.8	-6.8	-12.8	-18.8	-24.8	0.0
Male (This study)dB	-4.8	-2.0	0.0	-6.3	-10.8	-12.6	-15.2	0.0

Table 1. Octave levels (dB) relative to the A-weighted speech level from IEC 60268-16 and for the recommended levels from this study adjusted to produce 0 dBA.

4 INFLUENCE OF THE PROPOSED MALE SPEECH SPECTRUM ON THE STI

In this section, the influence of the proposed spectrum on STI calculations was investigated with seven sound systems and comparisons made with STI using the standard speech spectrum. The term “proposed male spectrum” refers to the data given in Table 1.

4.1 Seven sound systems.

Impulse response measurements made with six sound systems installed at public spaces and a virtual sound system designed for the stands of a typical medium size football stadium were employed for the STI calculations. All seven sound systems utilised distributed loudspeakers and were designed to produce speech levels above the background noise. Photos of the six public areas with their sound systems are shown in Figure 6 overleaf.

Acoustical parameters, SPL and STI, were assessed at twelve positions. The STIs were derived from the impulse responses using the Indirect Method⁶ (Schroeder equation) and post processing according to the methodology stated in IEC 60268-16:2011⁶



Conference room (1)



Grand Concourse at Sydney Central Station (2).



Suburban train station platform, in Sydney (3).



Overground Platform Central Station Sydney (4)



Underground station Sydney (5)



Road tunnel Brisbane Australia (6).

Figure 6. Photos of six of the seven public spaces employed for the STI calculations. The number stated for each area was used for future references in the text.

CATT Acoustic software V9.0c was used to predict the impulse responses and the system's output SPL for twelve positions at a virtual football stadium. Dimensions of the grandstands were taken from UEFA 2012²³ and absorption coefficients taken from Egan²⁴ 1988 and Kuttruff²⁵ 2009 were employed for the predictions. A scenario with 80% of the stands occupied with people was considered. Scattering coefficients were selected using the guidelines included in the CATT-Acoustic software manual. Temperature of 20° and humidity of 50% were also used for the predictions. A total of 26 loudspeaker clusters comprising of two JBL PD5322-95 loudspeakers were located in the model at intervals of 16 m on the canopy roof at 3.0 m from the front edge. Twelve receivers located at 1.2 m from the floor were distributed throughout the North stand area. Figure 7 shows the CATT model employed in the predictions.

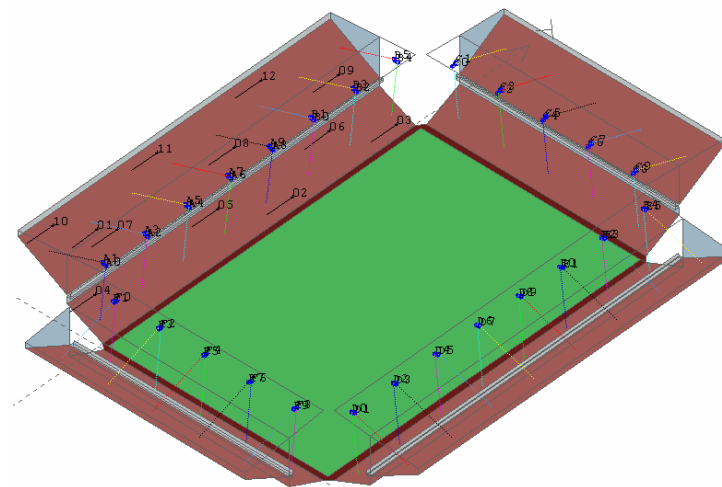


Figure 7. CATT model for the Football stadium (30,000 spectators).

4.2 Background Noise Levels

Crowd noise levels measured at the Etihad Football Stadium located in Melbourne, Australia, were used for the STI calculations. The Etihad stadium has a capacity for 53,359 spectators. The measured noise was extrapolated to the lower capacity of the stadium considered here (30,000 spectators), by reducing 2.5 dB the octave band levels, [$2.5 \text{ dB} = 10 \cdot \log(30,000/53,359)$].

Octave band L_{Eq} and percentiles background noise levels were also measured for each area. For the Road tunnel situation, the measured noise produced by smoke extractor fans distributed throughout the tunnel was employed for the STI calculations. Table 6 in the Appendix provides background noise levels for all the seven areas.

4.3 System Output Level

The output SPL of the systems was measured employing a standard pink-noise with the male spectrum recommended in IEC 60268-16:2011. Octave bands from 125 Hz to 8 kHz and overall A-weighted levels were obtained for each area at the selected positions by means of L_{Eq} measurements of 10 seconds duration. For consistency in the STI calculations, the SPLs were mathematically adjusted so that the arithmetic average level for each area produced a SNR of 10 dBA above the background noise for six the areas. The relative levels between the octave bands were kept constant in the process. For the underground station, a system level of 90 dBA was employed producing a SNR of 7 dBA. The SPLs employed in the STI calculations are given in Tables 6 and 7 in the Appendix.

4.4 STI calculations.

The following calculations were performed:

- 1) The STIs were calculated for each position with the system output levels included in Section 4.3 above and shown in Table 7, and the noise levels given in Table 6 (in Appendix).
- 2) The octave band levels at each measurement position were adjusted according to the proposed male spectrum given in Table 1.
- 3) The STIs were then calculated for each position with the adjusted octave band values and the background noise data of Table 6.

Table 2 shows the differences between the STI results obtained with the IEC spectrum and with the proposed spectrum.

Position	STI difference (Proposed male spectrum minus IEC values)						
	CR (1)	GC (2)	STS (3)	OP (4)	UP (5)	Rd. T (6)	FS
Pos 01	0.05	0.04	0.06	0.04	0.02	0.03	0.04
Pos 02	0.06	0.03	0.06	0.04	0.03	0.03	0.04
Pos 03	0.05	0.04	0.05	0.03	0.03	0.03	0.04
Pos 04	0.05	0.05	0.05	0.04	0.02	0.03	0.04
Pos 05	0.05	0.04	0.05	0.04	0.02	0.04	0.04
Pos 06	0.06	0.04	0.05	0.04	0.03	0.03	0.03
Pos 07	0.05	0.04	0.06	0.04	0.02	0.03	0.04
Pos 08	0.05	0.04	0.06	0.04	0.02	0.03	0.04
Pos 09	0.05	0.04	0.06	0.04	0.03	0.03	0.04
Pos 10	0.05	0.05	0.06	0.04	0.03	0.03	0.04
Pos 11	0.05	0.04	0.05	0.04	0.03	0.03	0.04
Pos 12	0.06	0.04	0.06	0.04	0.03	0.02	0.04

Table 2. STI difference between the calculation results with the IEC spectrum and the proposed male spectrum (seven sound systems). CR is the Conference room (1), GC is the Grand Concourse at Sydney central station (2), STS is the Suburban train station platform (3), OP is the Overground platform (4), UP is the Underground platform (5), Rd. T is the Road tunnel (6), and FS is the Football stadium.

An increase in the STI values with the proposed male spectrum can be seen in Table 2 for all receivers. The STI increases range from a minimum of 0.02 observed for positions of the Underground platform (UP) and the Road Tunnel (Rd.T), to a maximum of 0.06 obtained for positions of the Conference room (CR) and the Suburban train station platform (STS). The percentage distribution of the STI differences presented in Table 2 has been included in Figure 8.

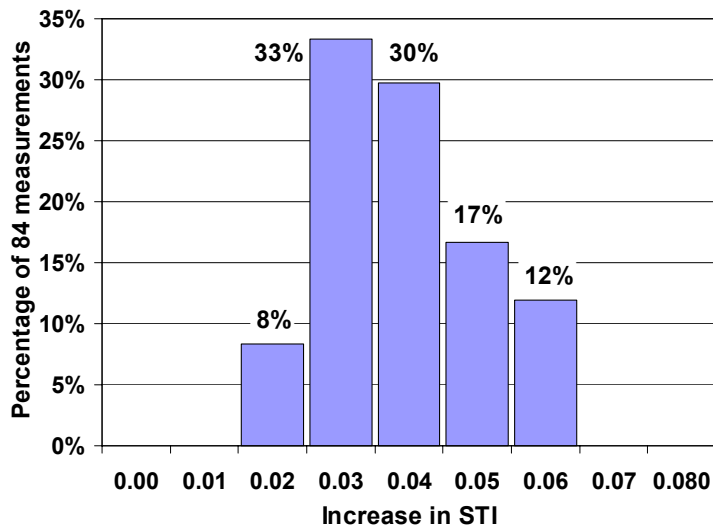


Figure 8. Percentage distribution of the increase in STI obtained with the proposed male spectrum for a total of 84 situations taken over the sound systems.

4.5 Discussion of Results

An STI increase of 0.02-0.06 was found for all the areas with the proposed male spectrum. These increases were due to the higher SNRs at mid and high frequencies with the proposed spectrum.

Table 3 relates to the Grand Concourse and shows the average octave-band SNRs and their differences calculated for the IEC and the proposed male spectrum that were set to identical A-weighted levels with 10 dB SNR.

Spectrum	Average signal-to-noise ratio (dB)						
	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz
1. IEC male	13.7	14.8	12.3	7.9	4.7	3.2	0.0
2. Proposed male	6.0	9.9	13.1	8.3	6.7	9.4	9.6
Difference (2-1)	-7.7	-4.9	0.8	0.4	2.0	6.2	9.6

Table 3. Averaged SNR calculated with the proposed male spectrum and with the IEC spectrum for the Grand concourse.

Table 3 shows that the use of the proposed male spectrum in the Grand concourse produced a decrease in SNR in the 125 Hz and 250 Hz octave bands, when compared with the IEC spectrum, which are the bands that contribute less to STI values. A small increase in SNR was found when using the proposed male spectrum at 500 Hz, 1 kHz and 2 kHz (2 dB or less), and more than 6 dB was found for the 4 kHz and 8 kHz bands. Similar differences in SNR between both spectra would be expected for all the remaining areas.

The changes in STI shown in Figure 8 depend on the SNRs at which the STIs were calculated. To illustrate this dependency, the STI was calculated at several SNRs and the STI change plotted against SNR. The system SPLs were kept constant and the background noise in each situation was adjusted in 1 dB increments whilst retaining its spectral shape. Figure 9 shows the results for two of the seven areas.

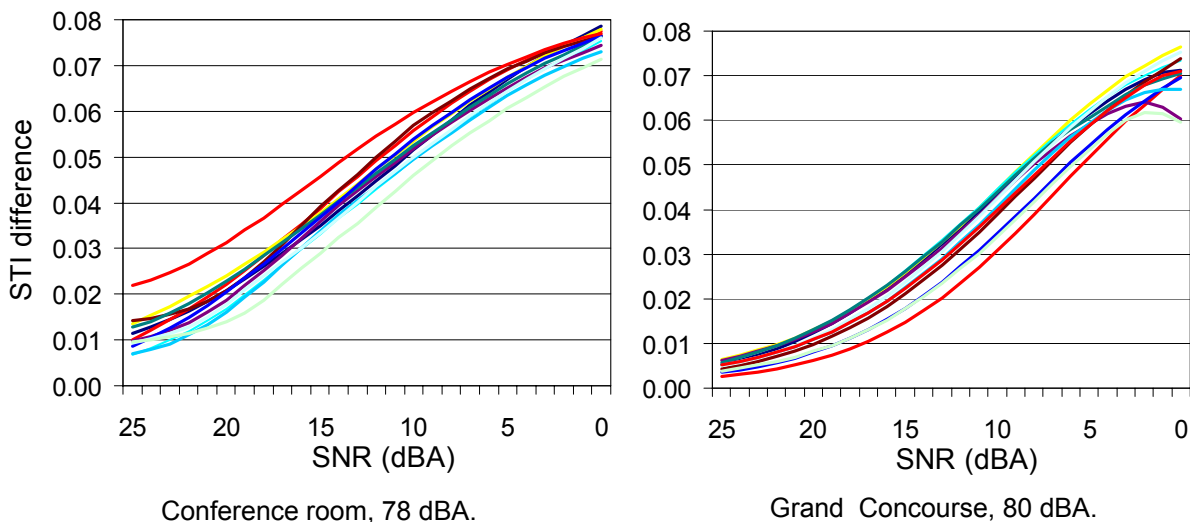


Figure 9. Differences in STI between IEC and the proposed spectra due to changes in SNR at 12 positions.

As expected, minimal STI differences were found for high (good) SNR conditions (0.01-0.02 for 25 dBA) whilst higher differences were found for low (poor) SNR conditions (0.06-0.08 for 0 dBA). The change in STI difference with SNR was found to be different between areas. Similar plots for remaining areas are given in Figure 11 of the Appendix.

Additional STI calculations with the proposed male spectrum were carried out for each area with lower average SPLs. The average SPL with the proposed male spectrum was iteratively reduced until the 5th percentile STI for each area matched the 5th percentile obtained with the IEC spectrum. Figure 10 shows the new average SPLs results.

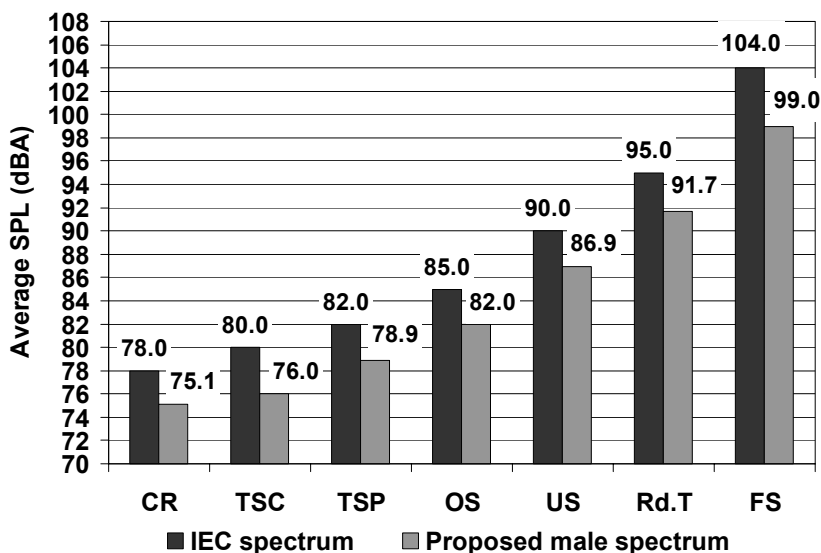


Figure 10. Average SPLs of sound system which are required with the proposed male spectrum to produce the same STI (5th percentile) as the IEC spectrum.

Figure 10 indicates that a lower system SPL for all the areas could be used with the proposed male spectrum, and still achieve a similar 5th percentile STI to the IEC spectrum. The decrease in SPL varies from a minimum of 2.9 dB for the Conference room to a maximum of 5.0 dB for the Football stadium.

4.6 Electrical power requirements.

The electrical power required by a sound system primarily depends on the required SPL and the spectrum of the input signal. Different input spectra require different electrical powers for similar SPL targets.

In order to ascertain the effect of the proposed male spectrum on the electrical power requirements, a measurement was carried out using two signals:

- a pink noise signal with the IEC male speech spectrum
- a pink noise signal with the proposed male spectrum.

These two signals were digitally band-pass filtered between 80 Hz and 12 kHz with 24 dB/octave filters and normalized to the same RMS level. Both signals matched the relative 1/3 octave bands target levels (100 Hz – 10 kHz) within 0.2 dB.

A-weighted voltage measurements of the normalised signals showed a 2.9 dBA higher level with the proposed male spectrum, compared with the IEC male speech spectrum for the same RMS signal level. This expected increase on the A-weighted SPL was tested at the 12 measurement positions in the Conference room, with the results being shown in Table 4.

	IEC speech	Proposed male spectrum
Average over 12 positions	75.0 dBA	77.8 dBA
Difference with IEC speech	-	2.8 dBA

Table 4. Averaged SPL in the Conference room with two input spectra: IEC male speech and the proposed male speech. The input signals were normalised to the same broadband RMS voltage.

The measured SPL increase of 2.8 dBA with the proposed speech spectrum reproduced by the Conference sound system agrees with the expected increase of 2.9 dBA over the IEC speech level. This increase would imply lower electrical power requirements with the proposed male spectrum to achieve similar SPL targets.

5 CONCLUSIONS

The following conclusions are made:

- In this study, the LTASS for male British English was measured for 40 people and found to differ substantially from the IEC 60268-16 spectrum specified for STI calculations.
- Based on the results, a new spectrum for male speakers is proposed for use with STI calculations.

- Examination of the literature indicated that inter-language differences are lower than inter-talker differences. This suggests that the proposed male spectrum could be applied to other languages without introducing large errors.
- The impact of the proposed male spectrum was assessed on the STI calculations and found to produce an increase of 0.02-0.06 for seven different types of sound systems operating under conditions of 10 dBA SNR.

Greater STI increases were found for lower SNR conditions whilst STI increases were minimal for quiet conditions.

- The greater high-frequency content of the proposed male spectrum compared with the IEC spectrum could allow a reduction in sound system SPL capacity for the same STI values as the IEC spectrum. Reductions in SPL between 3.0 dB and 5.0 dB were found with the seven sound systems investigated. These SPL reductions could be translated to reduced system cost and lower electrical power requirements.

6 FUTURE WORK

To enable the proposed male spectrum to be used with STI, further validation of the current STI for the English language would be required with the proposed male spectrum.

7. REFERENCES

1. ANSI S3.5-1986, "Method for calculation of the articulation index", American National Standards Institute, New York. (1986).
2. ANSI S3.5 (R2007), *American National Standard methods for the Calculation of the Speech Intelligibility Index*, (Acoustical Society of America, New York, 2007).
3. C.V. Pavlovic, "Derivation of primary parameters and procedures for use in speech intelligibility predictions," *J. Acoust. Soc. Am.* **82**, 413-422 (1987).
4. N.R. French and J.C. Steinberg, "Factors governing the intelligibility of speech of sounds," *J. Acoust. Soc. Am.* **19**, 90-119 (1947).
5. R.W. Benson and I. J. Hirsh, "Some variables in audio spectrometry", *J. Acoust. Soc. Am.* **25**, 499-505 (1953).
6. *IEC 60268-16-2011 "Sound system equipment. Part 16: Objective rating of speech intelligibility by speech transmission index*, (International Electrotechnical Commission, Geneva Switzerland).
7. S. S. Stevens, J.P. Egan and G. A. Miller, "Methods of measuring speech spectra," *J. Acoust. Soc. Am.* **19**, 771-780 (1947).
8. R.M. Cox, and J.N. Moore, "Composite speech spectrum for Hearing aid gain prescriptions," *J. Speech Hear. Res.* **31**, 102-107 (1988).
9. D. Byrne and H. Dillon, "The National Acoustic Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid," *Ear Hear.* **7**, 257-265 (1986).

10. A. Boothroyd, "The discrimination by partially hearing children of frequency distorted speech," *Int. Audiol.* **6**, 136-145 (1967).
11. T. Tarnoczy and G. Fant, "Some remarks on the average speech spectrum," Q. P.S. R. Rep. No. 4, pp. 13-14, Speech Transmission Laboratory, Stockholm, (1964).
12. D. Byrne, "The speech spectrum - Some aspects of its significance for hearing aid selection and evaluation," *Br. J. Audiol.* **11**, 40-46 (1977).
13. H. Kiukaanniemi, P. Soponen and P. Mattila, "Individual differences in the long-term speech spectrum," *Folia Phoniatr.* **34**, 21-28 (1982).
14. B. Harmegnies and A. Landercy, "Intra-speaker variability of the long term speech spectrum," *Speech Communication*, **7**, Issue 1, 81-86 (March 1988).
15. C.V. Pavlovic, M. Rossi, and R. Espesser, "Statistical distribution of speech for various languages," *J. Acoust. Soc. Am.* **88**, Issue S1, 176-176 (1990).
16. K. S. Pearsons, R.L. Bennett, and S. Fidell, "Speech levels in various noise environment," EPA Rep. No. 600/1-77-025 Environmental Protection Agency, Washington DC., (1977).
17. I.R. Cushing, F.F. Li, T.J. Cox, K. Worrall, and T. Jackson, "Vocal effort levels in anechoic conditions," *Applied Acoustics*, **72**, Issue 9, 695-701 (Sept 2011).
18. Byrne et al.: "Long-term average speech spectra. *J. Acoust. Soc. Am.* 96, No. 4, 2108-2120 (1994).
19. G. A. Studebaker, "Directivity of the human vocal source in the horizontal plane," *Ear Hear.* **6**, 315-319 (1985)..
20. W.T. Chu and A.C.C. Warnock, "Detailed directivity of sound fields around human talkers," Technical Report, Institute for Research in Construction (National Research Council of Canada, Ottawa ON, Canada), pp. 1-47 (2002).
21. IEEE Subcommittee on Subjective Measurements. IEEE Recommended Practices for Speech Quality Measurements. *IEEE Transactions on Audio and Electroacoustics*, Vol 17, 227-246, (1969).
22. P. Welch, "The use of the fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Trans. Audio Electroacoust.*, vol. AU-15, no. 2, pp.70-73 (Jun. 1967).
23. *UEFA Guide to quality Stadiums.* (UEFA 2012).
24. M. David Egan, *Architectural Acoustics*, (McGraw-Hill, Inc., New York, 1988).
25. H. Kuttruff, *Room Acoustics*, 5th ed. (Spon Press, Oxon, UK, 2009).

8. APPENDIX

Frequency	IEC 60268-16		Current study	
100 Hz	62.1	-	54.2	-
125 Hz	62.1	66.8	55.5	59.1
160 Hz	62.1	-	53.1	-
200 Hz	62.1	-	56.7	-
250 Hz	62.1	66.6	58.0	62.0
315 Hz	61.5	-	56.7	-
400 Hz	60.0	-	58.8	-
500 Hz	58.4	63.2	60.3	63.9
630 Hz	56.5	-	58.0	-
800 Hz	54.4	-	52.0	-
1 k	52.4	57.4	53.6	57.7
1.25 k	50.4	-	52.9	-
1.6 k	48.4	-	50.2	-
2 k	46.4	51.4	46.9	53.2
2.5 k	44.4	-	47.3	-
3.15 k	42.4	-	47.0	-
4 k	40.4	45.4	46.8	51.4
5 k	38.4	-	45.9	-
6.3 k	36.4	-	45.3	-
8 k	34.4	39.4	44.1	48.7
10 k	32.5	-	41.6	-
dBA	64.0		64.0	

Table 5. SPL (dB) for the IEC spectrum and for the male spectrum proposed in the current study. All levels were normalised to an overall SPL of 64 dBA. Integrated octave band values are also given.

Area	Background noise (L10,dB)							
	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz	A-w
Conference room	75	69	64	61	59	57	55	68
Grand Concourse	69	68	67	65	62	58	55	70
Suburban train station	68	68	71	67	62	59	59	72
Overground station	72	75	74	69	65	63	57	75
Underground station	78	81	83	77	72	65	58	83
Road tunnel	80	85	82	80	76	70	62	85
Football stadium	76	81	91	91	86	77	65	94

Table 6. L10 background noise levels measured at six public areas and measured noise levels at the Etihad stadium reduced by 2.5 dB.

Area	Octave band levels with IEC input spectrum (dB)							
	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz	dBA
Conference room	82	81	76	71	66	60	58	78
Grand concourse	83	83	79	73	67	62	55	80
Suburban train station	76	82	82	77	70	62	55	82
Overground station	81	87	84	79	73	67	60	85
Underground station	87	93	89	84	79	72	63	90
Road tunnel	50	91	95	89	83	75	68	95
Football stadium	107	107	103	97	91	85	79	104

Table 7. Octave band and A-weighted SPL employed for the STI calculations. The averages for 12 positions are given for each area.

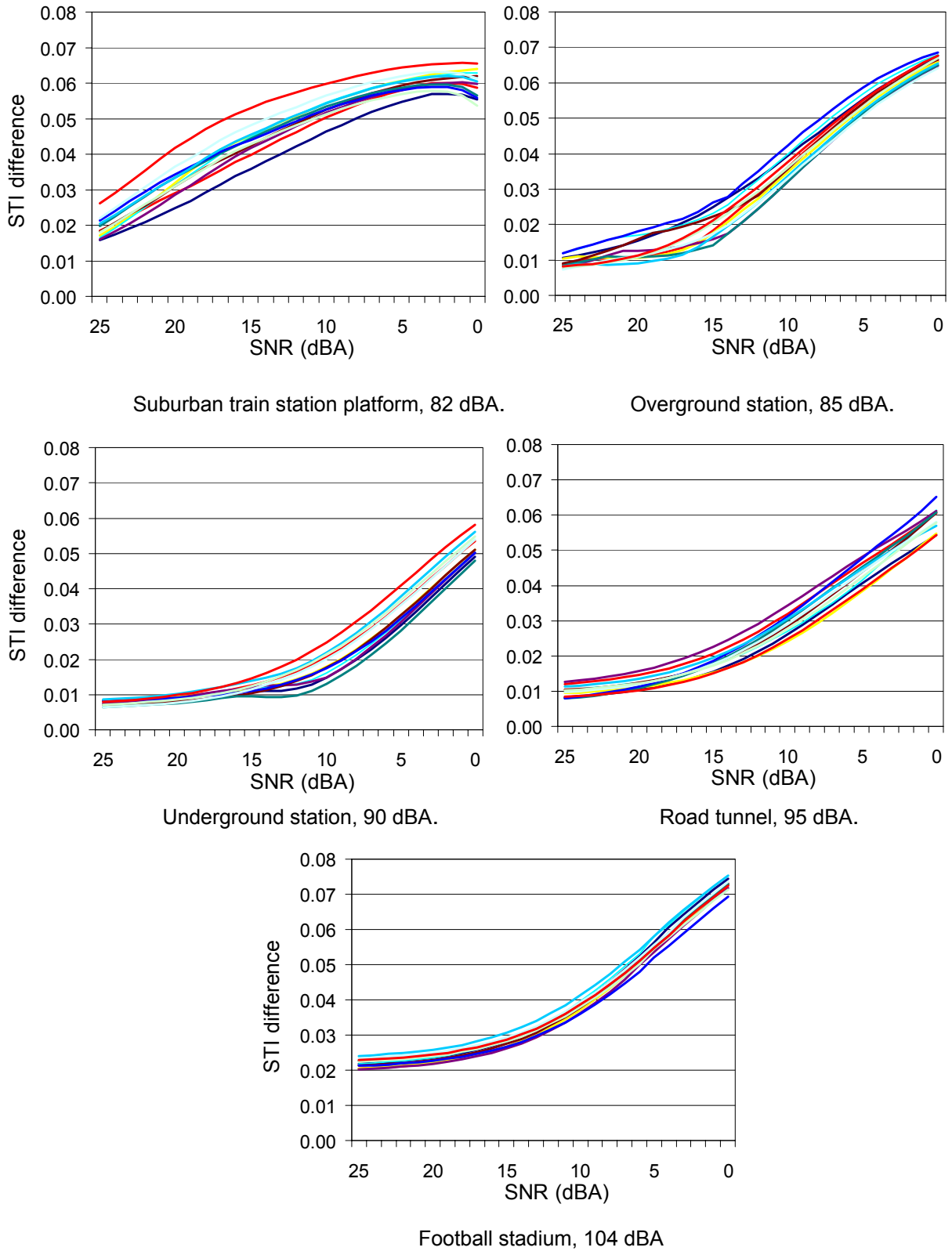


Figure 11. Influence of the SNR on the difference in STI found between the use of the IEC spectrum and the proposed male spectrum. The SPL of the sound system has been indicated for each area.