THE EFFECT ON STI RESULTS OF CHANGES TO THE MALE TEST-SIGNAL SPECTRUM

Glenn Leembruggen	Acoustic Directions, ICE Design, University of Sydney, Australia
Jan Verhave	Embedded Acoustics, Delft, The Netherlands
Stefan Feistel	AFMG, Berlin, Germany
Ludger Holtzem	NTi Audio, Schaan, Liechenstein
Peter Mapp	Peter Mapp Associates, Colchester, UK
Hiroshi Sato	AIST. Ibaraki, Japan
Thomas Steinbrecher	Bose, Friedrichsdorf, Germany
Sander Van Wijngaarden	Embedded Acoustics, Delft, The Netherlands

1 INTRODUCTION

During recent maintenance work on the STI standard IEC 60268-16 (1), several questions arose concerning the spectrum of the male test signal used in the STI method. With Edition 4 and earlier editions of the standard, the spectrum of the male test signal is formed from a pink noise signal combined with a 6 dB/octave roll-off from 500 Hz upwards. This spectrum shows important differences from male speech spectra as found in literature (2) (3) (4), in particular in its comparatively strong low frequency content.

In addition, from a practical point of view, the male speech spectrum of Edition 4 has some electro-acoustic drawbacks due to its relative strong low-frequency content.

Electro-acoustical equipment sized to a small head sized (e.g. artirficial mouths) will not easily meet the requirements imposed by the STI test signa, which are low distortion and head-like directivity. Low directivity up to the 8kHz octave calls for small-aperture drivers, which must be also capable of reproducing high level low-frequency content up to 90 dB peak at 1 m without significant distortion. (Raised voice of 75 dB with a crest factor of 15 dB)

Even though a few selected products comply with these requirements, many high-end products are not capable of generating sound pressure levels higher than the nominal speech level of 60dBA at 1 meter without significant harmonic distortion and buzz.

This limitation forced the standardisation committee of NFPA 1981:2013 to reduce the sound pressure levels for the 125 Hz and the 250 Hz octave band by 9 and 2 dB respectively to enable reproduction of a raised voice level up to 75 dBA at 1 meter for assessing face masks using the Kemar head and torso simulator.

With the above-mentioned issues in mind, the STI Maintenance Team (MT) decided to investigate the possibilities of changing the male test-signal spectrum in such a way that it would be more similar to published spectra whilst having minimal impact on STI values. The team agreed that a reduction of 6 dB in the 125 Hz octave band and a 3 dB reduction in the 250 Hz octave band would be appropriate.

An investigation was conducted using Matlab software to determine the extent to which STI values would be affected by a range of spectral changes.

2 HISTORICAL BACKGROUND ON STI TEST SIGNALS

2.1 Research Before 1974

The spectrum of the current STI test signal is primarily based on research undertaken during the 1980's and early 1990s at TNO in the Netherlands. TNO's report (5) describes the recording and analysis of different speech material with four males and four female talkers:

- CVC words embedded in a carrier sentence (51 different CVC words in 5 different carrier sentences)
- SRT sentences (26 sentences, approximately 6 words each)
- Connected discourse (full text, approximately 264 words in length, normal and whispered)

The recordings were made in an anechoic room, using a blanket-covered table and a B&K 4155 recording microphone. The distance of the talker and the microphone was approximately 0.5 meter and at a 15º angle from the centre axis. Recordings were stored on DAT tape and spectral analysis was done with a Rohde and Schwarz 1/3 octave band analyser (type FAR). Measurements (100 msec frames, fast weighing) were averaged over 60 seconds. The octave band spectra relative to an overall A-weighted level of 0 dBA of male speech material are shown in Table 1 and Figure 1.

Octave band	Embedded CVC words	Sentences	Connected discourse	Average
125	1.8	0.7	2.1	1.6
250	3.2	2.7	4.0	3.3
500	-0.2	-0.2	-1.8	-0.7
1000	-9.1	-8.8	-7.2	-8.3
2000	-12.4	-11.5	-12.3	-12.0
4000	-17.9	-18.2	-15.6	-16.5
8000	-22.6	-21.3	-22.1	-22.0

Table 1 Male speech spectra in dB relative to 0 dBA as a function of octave band frequency for TNO speech material.

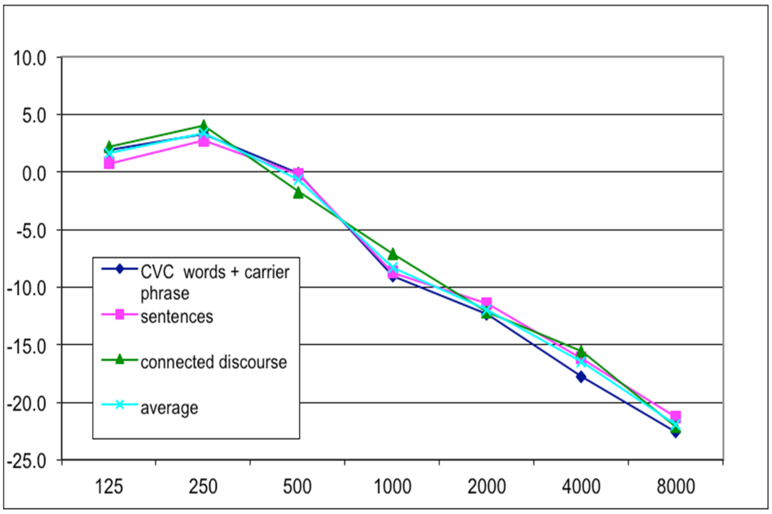

Figure 1 Speech spectra in dB relative to 0 dBA as a function of octave band frequency for TNO speech material.

Between 1989 and 1991, Steeneken and Van Velden performed further research on the spectra of different phoneme groups. For that purpose, speech material of 10 male and 10 female talkers (stressed and unstressed) was recorded. Due to a relatively low sample rate (12820 Hz), the highest octave band was 4 kHz. Results of that research is shown in Fig. A6.1 on p159 of (6). Figure 1 shows the phonetically balanced spectra produced from that work.

2.2 Research After 1993

Table 2 shows speech spectra derived from (2), (3), and (4). Prior to 1993, the STI test signal was produced by hardware known as STIDAS, which used resistor/capacitor networks to shape the spectrum. In 1993 the male STI test signal was generated using an analogue pink noise generator in combination with -6 dB/octave shaping filter from 500 Hz and up. Evaluation of the revised STI version (STIr, 1997) was also done using this first order approximation of the male speech spectrum. The spectrum obtained with the STIDAS networks was recently re-measured and is shown in Table 2 (STIDAS R//R).

The spectrum titled "Mean LABS" shows the arithmetic mean in dB in each octave band of spectra obtained by Byrne (2) Cushing (3), ANSI (4), and Steeneken (6). Additionally, the current STI spectrum (STI 2011) and a candidate for the new STI spectrum "IEC 2016". All speech spectra are scaled to an A-weighted level of approximately 70 dBA.

Figure 2 shows graphically the data of Table 2.

Octave band	Cushing	ANSI	Byrne et al	Steeneken PB	Mean LABS	STIDAS R//R	IEC 2011	"IEC 2016"
125	61.7	58.9	65.3	71	65.1	71.7	72.8	67.3
250	68.4	67.6	68.7	72	69.4	71.9	72.8	70.3
500	69.4	69.9	70.5	69	69.8	69.2	69.1	69.6
1000	64.3	64.8	63.6	63	63.8	64.1	63.1	63.6
2000	60.2	59.5	58.6	58	58.7	58.4	57.1	57.6
4000	56.2	54.3	54.4	55	54.5	51.1	51.1	51.6
8000	50.7	47.8	52.2	-	50.0	45.0	45.1	45.6

Table 2 Speech spectra and STI test signal spectra in dB in from different laboratories as a function of octave band frequency. All spectra are normalised to 70 dBA.

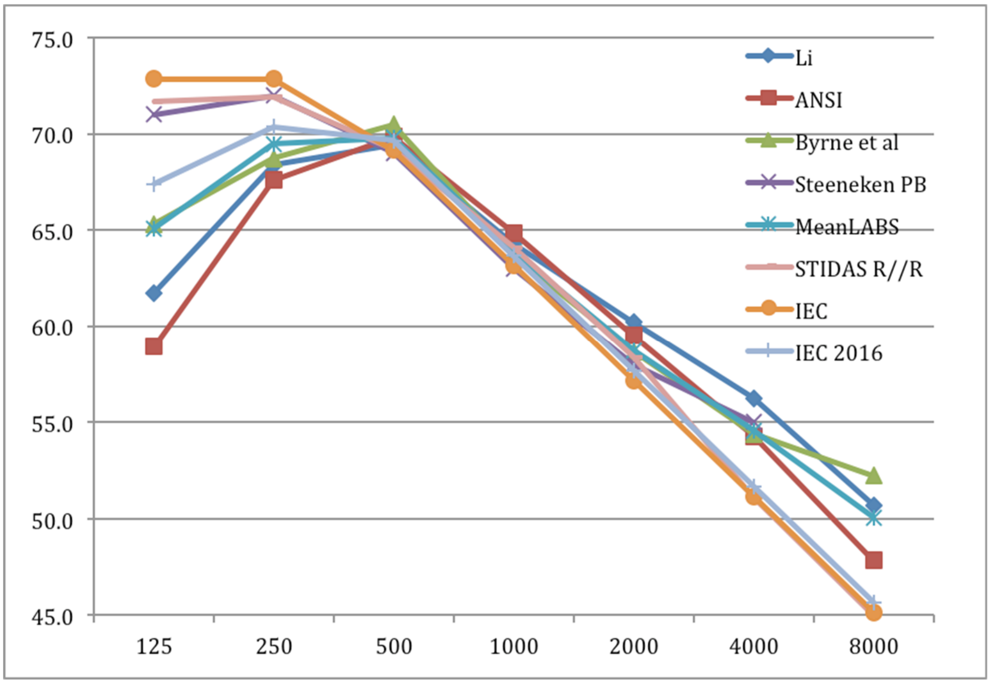

Figure 2 Speech spectra and STI test signal spectra from different laboratories in dB as a function of octave band frequency

3 ERROR SIMULATION

3.1 STI in Noise Using Different Test Signal Spectra

It is likely that changes to the test-signal spectrum would introduce changes to the STI, due to the different signal-to-noise ratios in octave bands. A deviation in the STI would be acceptable if it remains within a specific limit. A recognized and acceptable deviation is 0.03 STI, which can be directly related to a change of approximately 1 dB in the effective signal-to-noise ratio. However, such a deviation is only acceptable if the STI errors are not positively or negatively biased and the mean of error distributions remains close to zero.

Simulations were carried out to investigate the range of errors that would occur with changes to the test-signal spectrum.

3.2 Simulation Parameters for Random Noise Spectra

The STI has two major dependencies with respect to the level of the test signal:

- the overall sound pressure level, which affects the level-dependent masking parameter
- the signal to noise ratio (SNR) in each octave band.

To ensure that simulations covered the entire range of STI values of 0 to 1, overall sound pressure levels ranging from 20 dBA SPL to 120 dBA SPL and SNR ranging from -12 to +43 were used.

For each speech spectrum, 250 random noise spectral shapes were generated. Each noise spectrum was then added to the speech spectrum based on an arbitrary SNR (-12 to +43) in both STI signals (i.e. the original and modified spectra). The total level of the noise and STI signal was then set at a level between 20 and 120 dB SPL.

For each condition, the original STI and the “modified” STI (reflecting the new spectrum) STI were calculated. This resulted in 1,414,000 (250 x 101 x 56) STI value pairs per speech spectrum. Using this value pairs (which were equally distributed for the full STI range), the percentage error was calculated for specific error thresholds.

The worst-case noise situation for STI occurs when the noise has the spectrum of male speech, as the signal-to-noise ratios (SNR) are the same in each octave band.

The interfering noise spectra could have been chosen more specifically since the only octave bands affected by the change in spectra are the 125, 250 and 500 Hz[1]. It was however decided to vary the noise spectrum within all the octave bands.

Figure 3 shows the range of errors in STI between the modified spectra and current spectrum as a function of the STI for a typical example. In this example, one random noise spectrum is added to a selected male test-signal spectrum at different signal levels (101 steps of 1 dB) and at 56 different signal-to-noise ratios (SNR) with steps of 1 dB. Some bias is evident, as the error values are not normally distributed; however, both negative and positive error values do occur. It is interesting that in this example, the exceedances of 0,03 STI occur mainly in the region of 0.6 - 0.8 and generally not for low STI values.

Figure 4 shows the distribution (histogram) of the STI values with a bin size 0.02 STI and indicates that the values are essentially equally-distributed with approximately 2% per 0.02 STI bin.

[1] 500 Hz is included here as the m values in this band include masking by the 250 Hz octave band,

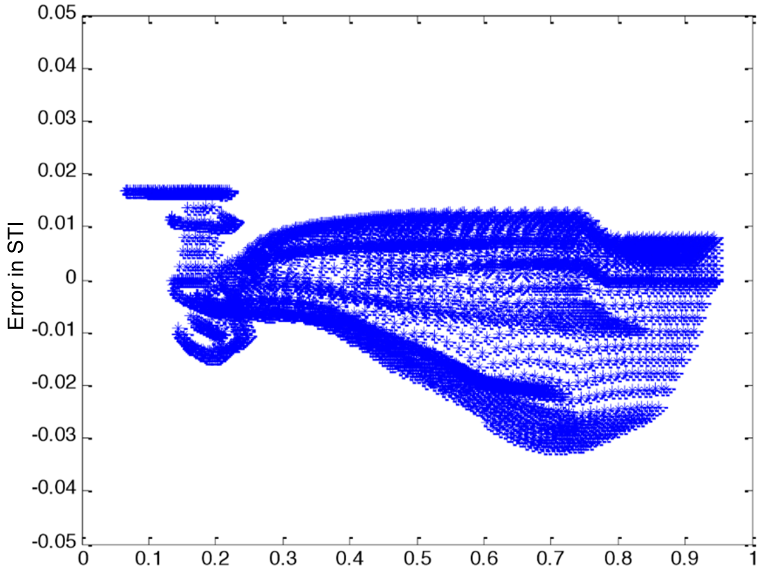

Figure 3. Graph of typical error values between the STIs of a selected spectrum and the existing spectrum as a function of the STI using the current spectrum.

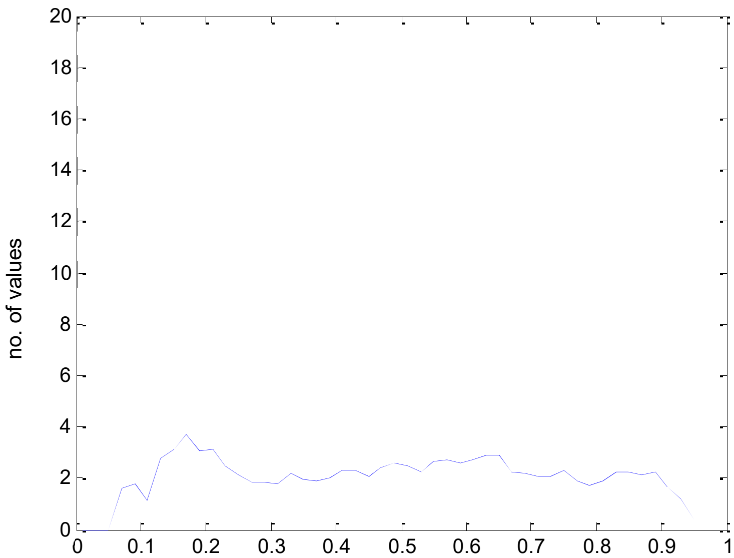

Figure 4. Distribution of the number of STI values per STI 0.02 STI range.

3.3 Range of Speech Spectral Changes Examined

A number of changes to the male speech spectrum were selected for investigation, and are shown in Table 3. The STI error values were the calculated for these selections. The proposed change for the forthcoming revision of the STI standard is Spectrum S4.

Spectrum	S1	S2	S3	S4	S5	S6	S7	S8	S9
125 Hz	-3	-6	-9	-6	-9	-12	-9	-12	-15
250 Hz	0	0	0	-3	-3	-3	-6	-6	-6

Table 3 Selected changes to the speech spectra examined relative the current male spectrum

It is worth noting that these reductions can produce an intrinsic reduction in the MTI value in the 125 Hz band under perfect transmission conditions. This results from the reduced signal-to-noise ratio of the test signal and the noise representing the reception-threshold (viz. hearing threshold), which is 46 dB in the 125 Hz band.

For example, with speech at the nominal level of 60 dBA, the 6 dB reduction in the 125 Hz band produces a MTI value of 0.94 rather than 1 with the original spectrum, due to the reduced SNR of the test signal.

When the overall level of the test signal is scaled to a given level in dBA, the levels in the octave bands 500 Hz to 8 kHz will be slightly higher than with the current spectrum. These higher levels will provide higher SNRs, leading to systematically higher STI results with the proposed spectrum. With the proposed spectrum, the bias in STI results would be 0.015 STI, due higher SNR of 0.5 dB in the 500 Hz to 8 kHz bands.

Figure 5 compares the STI values of the proposed and existing speech spectra obtained with perfect transmission conditions, (i.e noise-free and anechoic conditions) as a function of the A-weighted speech level. The upper chart gives the actual STI values, and the lower chart gives the change in STI values.

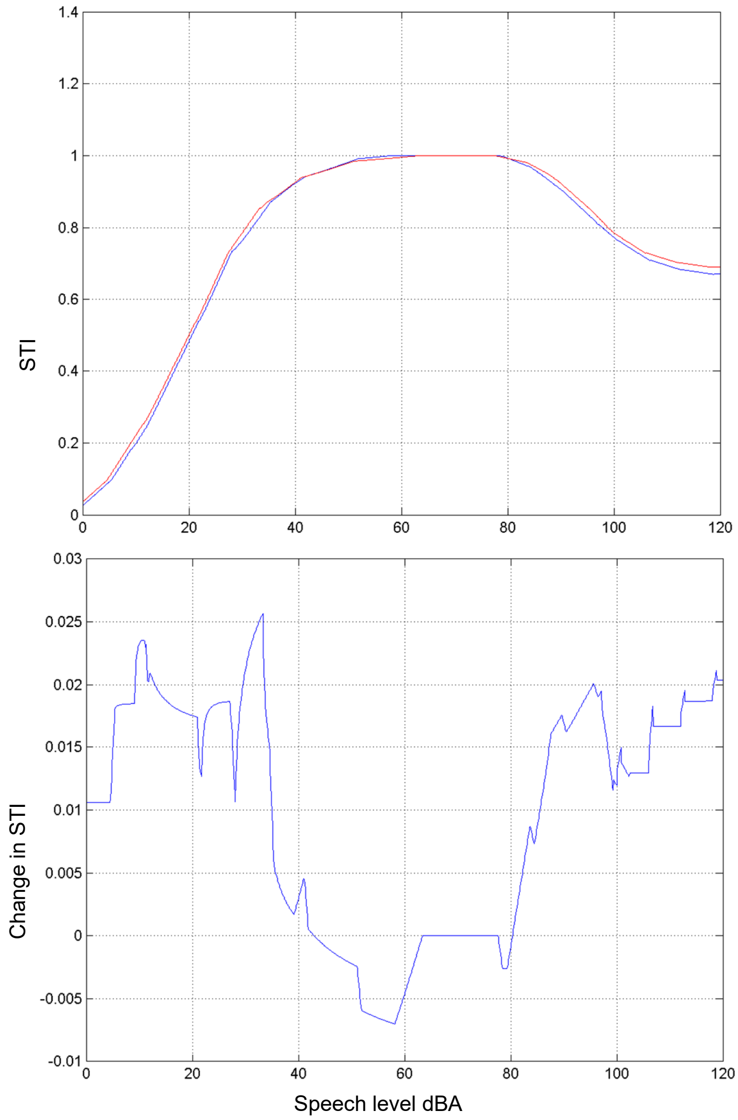

Figure 5 Comparison of STI values with proposed and existing speech spectra under perfect transmission conditions as a function of A weighted speech level.

3.4 Results with Randomly Generated Noise Spectra

For each of the spectral changes listed in Table 3, the percentages of STI error values exceeding a specific absolute STI threshold value were calculated by applying 250 random noise spectra, with the 101 steps in overall level and the 56 SNRs (approximately 1.4 million data values). Figure 6 shows the percentages exceeding each error threshold over the full STI range of 0 to 1. (N.B The graph's legend shows the values in the 125 Hz and 250 Hz bands as [125 250]). It is noted that the graphs show the maximum absolute error values, and the RMS values (standard deviation) of the errors are significantly lower.

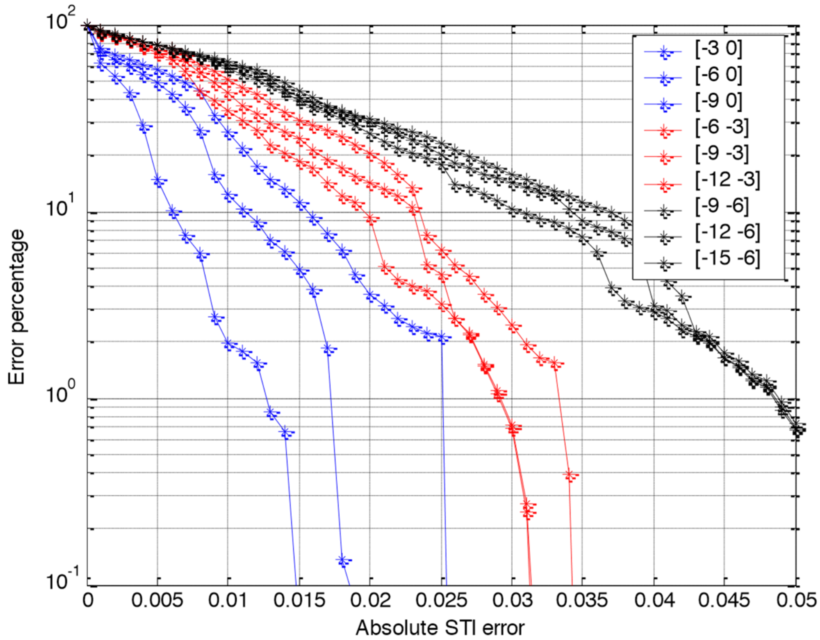

Figure 6 Percentages of STI error values as function of absolute STI error threshold for the full STI range 0 to 1.

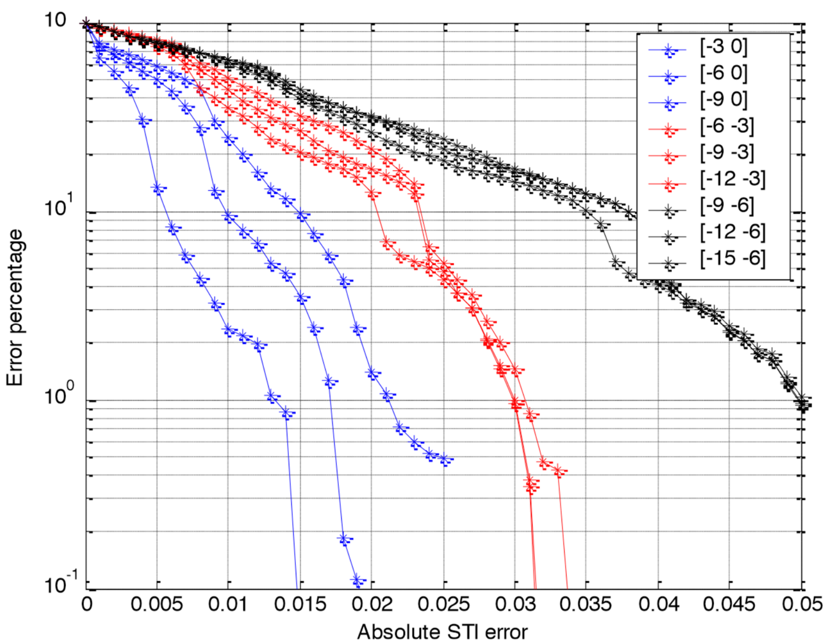

Figure 7 Percentages of STI error values as function of absolute STI error threshold for the STI range of 0.2 to 0.9.

Figure 7 presents results using a restricted and more realistic STI range of 0.2 to 0.9. In this figure, the bulk of the errors result from changes in the 250 Hz band.

For example, with a-6 dB level change for 250 Hz (black curves), the differences between the modified and original spectra exceed the absolute STI difference of 0.03 for at least 10 % of the situations. Also, for spectra with 0 dB and -3 dB in the 250 Hz band, less than 2 % of the error values are greater than 0.03.

Discussion

Figure 7 shows that the combination of a realistic STI range and random noise, a spectral change of up to -12 dB for 125 Hz and -3 dB for 250 Hz can be applied without a significant error in STI.

Table 4 and 5 summarise the statistics for each spectrum and overall range of STI values. The mean error-values are close to 0 indicating a very small and negligible bias in STI.

Spectrum	Mean	Std Dev	Min	Max
S1	0.0002	0.0039	-0.0192	0.0098
S2	0.0007	0.0069	-0.0224	0.0174
S3	0.0015	0.0092	-0.0245	0.0270
S4	0.0001	0.0113	-0.0336	0.0187
S5	0.0006	0.0127	-0.0355	0.0264
S6	0.0018	0.0143	-0.0366	0.0350
S7	0.0017	0.0182	-0.0615	0.0301
S8	0.0021	0.0194	-0.0610	0.0354
S9	0.0028	0.0202	-0.0600	0.0431

Table 4 Mean, standard deviation, minimum and maximum error values for the full STI range.

Spectrum	Mean	Std Dev	Min	Max
S1	-0.0004	0.0039	-0.0192	0.0094
S2	-0.0005	0.0066	-0.0224	0.0171
S3	0.0001	0.0088	-0.0245	0.0255
S4	-0.0021	0.0118	-0.0336	0.0177
S5	-0.0019	0.0131	-0.0355	0.0257
S6	-0.0011	0.0144	-0.0366	0.0346
S7	-0.0019	0.0192	-0.0615	0.0301
S8	-0.0017	0.0204	-0.0610	0.0354
S9	-0.0013	0.0207	-0.0600	0.0427

Table 5 Mean, standard deviation, minimum and maximum error values for a limited STI range of 0.2 to 0.9.

3.5 Results with RC and NCB Generated Noise Profiles

The range of errors in STI between the various spectra and the current spectrum was also explored with two commonly used noise spectrum profiles; RC (room criteria) and NCB (balanced noise criteria).

There are 26 RC profiles and 13 NCB profiles, each with a specific noise level in each octave band.

To increase the number of STI error results by using more SNRs, each profile was varied in 56 steps of overall level, whilst retaining the relative levels between octave bands. The total number of data pairs with the RC curves, was 147,056 (26 x 101 x 56), while the total number of data pairs with the NCB curves, was 73,528 (13 x 101 x 56).

Figures 8 and 9 show results for the RC and NCB profiles respectively over the restricted STI range (0.2 to 0.9). Both figures show that with spectrum S4, the percentage of STI values with a difference of 0.03 STI between the current and proposed speech spectrum is very low at less than 0.1 %.

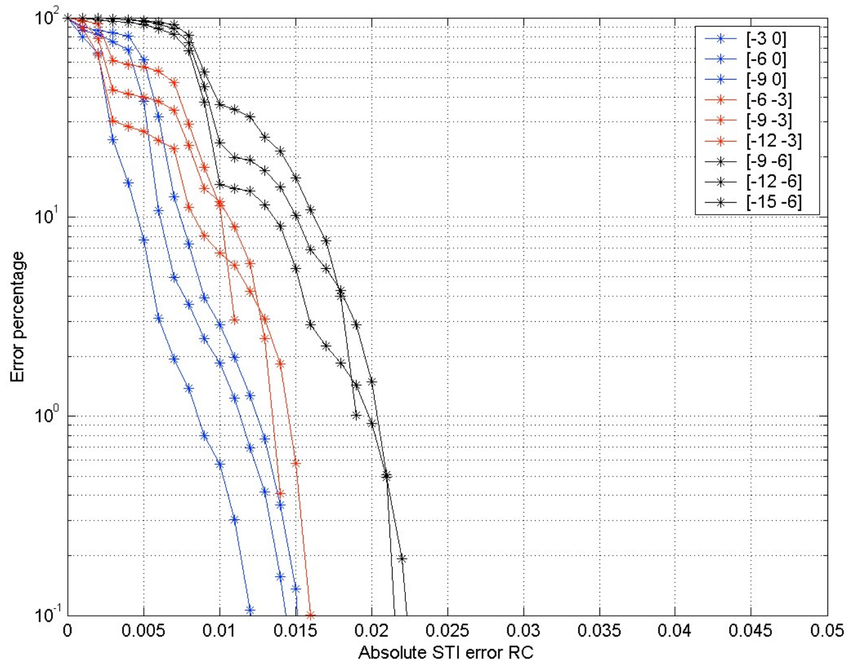

Figure 8 Percentages of error values exceeding an absolute threshold with profiles RC noise profiles for a limited STI range.

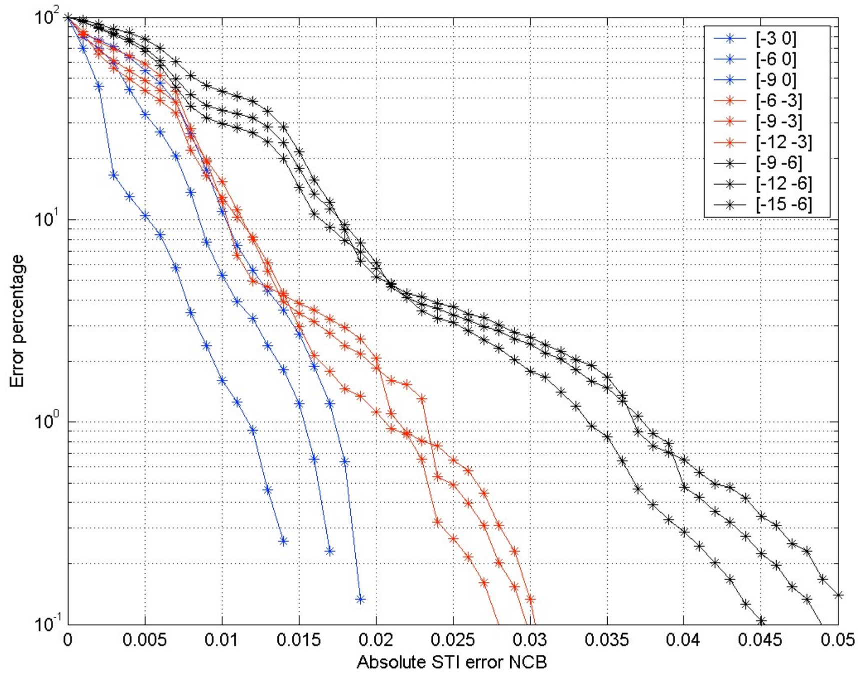

Figure 9 Percentages of error values exceeding an absolute threshold with profiles NCB noise profiles for a limited STI range.

4 CONCLUSION

The simulations of the changes to STI from the use of alternative speech spectra have been conducted using Matlab of more than 1.6 million scenarios. The results indicate that the selection of a 6 dB reduction for the 125 Hz octave band and a 3 dB reduction for the 250 Hz band is without doubt defendable. The difference (or error) in the mean STI values between the proposed and current spectra will be less than 0.01 STI. The standard deviation of the errors is less than 0.015 STI (equivalent to 0.5 dB SNR), while the maximum error is 0.03 STI (equivalent to 1 dB SNR). Only in 2 % of the cases will the errors exceed the accepted measurement uncertainty level of 0.03 STI. For these cases, the difference is less than 0.035 STI.

In combination with RC and NCB noise profiles, the percentage values exceeding 0,03 STI is even smaller and less than 0.1 % for the -6 and -3 dB reduction for 125 Hz and 250 Hz.

The changes to STI values resulting from the proposed spectrum will be extremely small and generally well within the accepted STI measurement uncertainty. Given the benefit of reduced electro-acoustic power in the 125 Hz and 250 Hz and the fact that the proposed spectrum more accurately reflects contemporary male-speech spectra, the proposed spectrum is regarded as a suitable modification to the STI method.

5 ACKNOWLEDGEMENT

The other authors are indebted to co-author Jan Verhave for his mathematical and Matlab work for this paper and his invaluable insights around STI issues.

6 REFERENCES

1. Sound sytem equipment: Objective rating of speech intelligibility by speech transmission index: IEC, 2011. IEC 60268 - Part 16.
2. Byrne, et al. An international comparison of long-term average speech spectra. *JASA,.* 1994, Vol. 96, 4.
3. Cushing, Li, Cox, Worral, Jackson. Vocal effort levels in anechoic conditions. *Applied Acoustics.* 2011, Vol. 72, pp. 695-701.
4. ANSI. Methods for Calculation of the Speech Intelligibility Index. *American National Standard.* S3.5-1997.
5. Steeneken, Geurtsen and Agterhuis. *IZF 1990 A-13.*
6. Steeneken, H. *On measuring and predicting speech intelligibility.* s.l. : http://publications.tno.nl/publication/34610762/AF5jDY/steeneken-1992-on.pdf, 1992.