

THE EFFECT OF SOUND SOURCE DIRECTION ON SPEECH INTELLIGIBILITY IN ROOMS

S Dance Acoustics Group, London South Bank University, UK
S Toone Acoustics Group, London South Bank University, UK
J. Sheng Tham Acoustics Group, London South Bank University, UK
P. Mapp Peter Mapp Associates, UK

1 INTRODUCTION

The topic of speech intelligibility is still being discussed and researched as evidenced by the latest iteration of the British Standard EN IEC 60268-16: Sound system equipment - Objective rating of speech intelligibility by speech transmission index [1]. Speech intelligibility is complex and multifaceted with EN ISO 9921:2003. Ergonomics – Assessment of speech communication' serving as a guide to ensure adequate speech intelligibility in public spaces [2]. One area that has rarely been studied is human subject-based experiments and corresponding objective electroacoustic measurements. In particular, when the subject is not facing the sound source. For instance, when standing on an underground station, the loudspeaker is typically positioned behind the commuter. As such, experiments were undertaken in the acoustic laboratory of London South Bank University to understand this effect on both native and non-native English speakers

2 EXPERIMENTAL METHODOLOGY

The experimental methodology was developed to explore electroacoustic vs subject-based-speech intelligibility assessment under a variety of conditions. The electroacoustic measurement experiment used STIPA to assess the potential intelligibility using a calibrated class 1 Nor140 sound level meter (omni-directional microphone), with the test signal reproduced via Yamaha HS50 loudspeaker producing 65 dBA at the receiver position. A pink noise interference signal was produced by a NTI Minirator Pro connected to another Yamaha HS50 loudspeaker (57 dBA at the receiver). The sound levels were measured 3.3m from the sound source at a height of 1.2m, at the centre of the listening panel, see Figure 1. The PB wordlists were played via a Digigram VX pocket sound card connected to the Yamaha HS50 source loudspeaker and also set to produce an average (Leq) level of 65 dBA again at the receiver position. The speech, STIPA and ambient noise spectra $L_{Zeq, 20 \text{ seconds}}$ is given in Figure 2. The ambient noise was measured to be 36 dBA. Two STIPA measurements were taken each time to ascertain if the result was valid. The experiment was conducted using three room conditions in a reverberation chamber (200m³) with eight individuals present: (1) empty, (2) with additional 10m² of Class A porous absorption and (3) with an additional 20m² of Class A porous absorption. Reverberation time, T_{20} , was measured using the impulsive method (Balloon) using the Nor140 sound level meter.

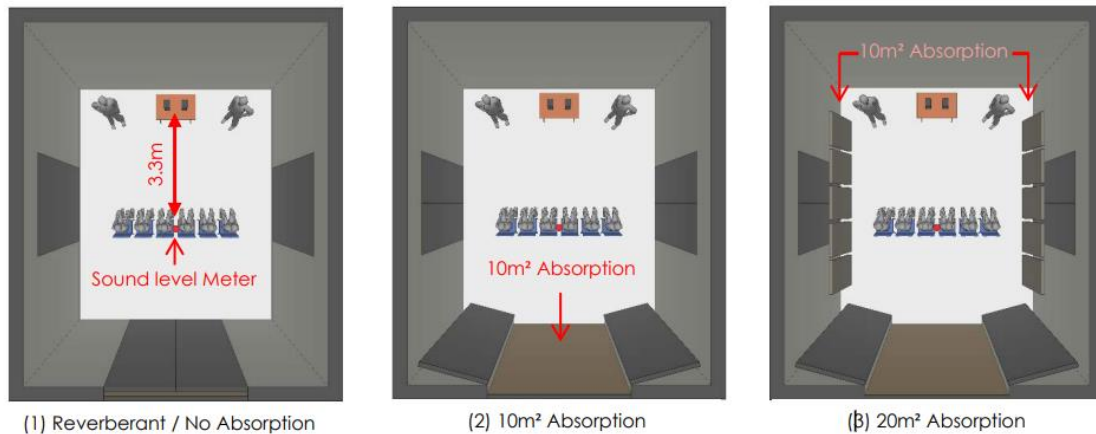


Figure 1. Three test conditions of the reverberation chamber and measurement setup

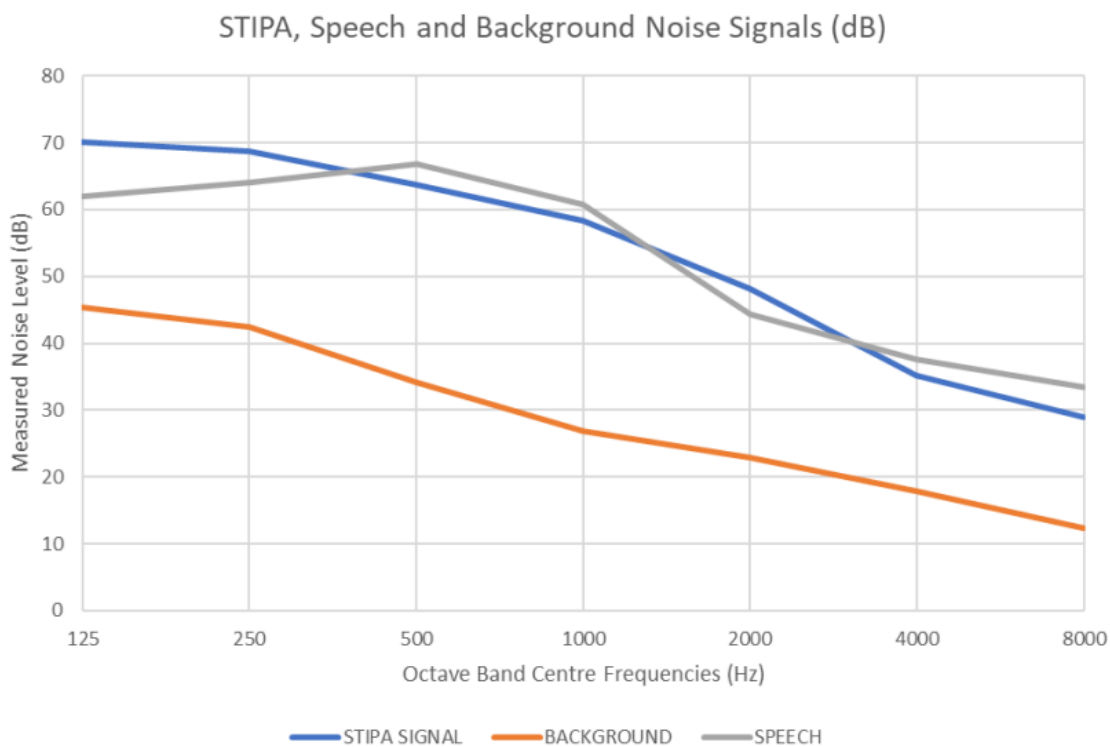


Figure 2. The speech, STIPA and background noise condition, $L_{Zeq, 20 \text{ seconds}}$ (dB)

The Subject based tests were undertaken by six seated students (all reported good hearing acuity): 3 native English speakers, and 3 non-native English speakers. Phonetically balanced (PB) words were used, reproducing the work of Morales [3-4]. In total, four 50 words lists were used, using normalised speech levels, all male voices. The subject based tests were undertaken under two room conditions: empty and with an additional 10m² of sound absorption employing the two previous signal-to-noise ratios, >15 dB and 8 dB. Finally, the test was retaken with the listening panel-facing away from the loudspeakers, see figure 3. All subjective and electroacoustic based tests are summarised in table 1.



Figure 3. Subjective testing: facing and facing away from the sound source, absorptive test condition.

Table 1. Electroacoustic and human subject test configurations

Test	Objective	Test	Subjective
Empty reverberation chamber with STI signal only	Yes	Empty reverberation chamber	Yes
Empty reverberation chamber with STI and noise signal	Yes	Empty reverberation chamber with noise signal	Yes
10m2 absorption with STI signal only	Yes	10m2 absorption	Yes
10m2 absorption with STI and noise signal	Yes	10m2 absorption with noise signal	Yes
20m2 absorption with STI signal only	Yes	10m2 absorption (reverse)	Yes
20m2 absorption with STI and noise signal	Yes	10m2 absorption with noise signal (reverse)	Yes

3 RESULTS

Reverberation time, T_{20} , was measured using an impulsive method, for three conditions. The T_{mf} (500, 1000, Hz averaged) values were as follows: empty 2.85 s, 10m² 1.43 s and 20m² 0.85 s, see figure 4. All measurements were taken with eight individuals in the room and six chairs.

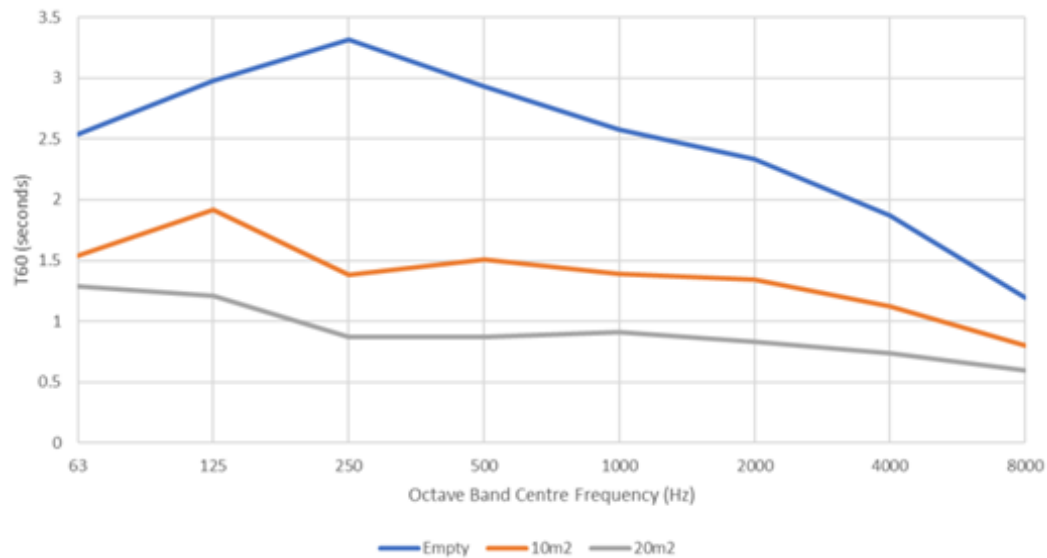


Figure 4. Reverberation time, T_{20} , measured under three configurations.

The STIPA measurements are presented in figure 5 under six conditions.

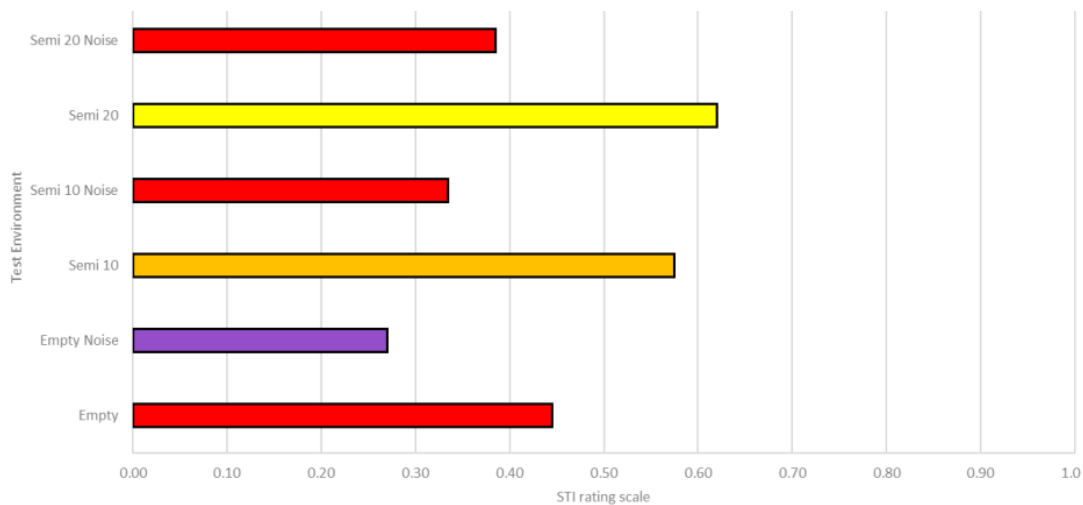


Figure 5 shows the STIPA results for the three room configurations under two noise conditions.

The electroacoustic STI speech intelligibility test, presented in Figure 5, showed that reducing the signal-to-noise ratio from >15 dB to 8 dB resulted in a 0.15 drop in STIPA values under the three room configurations. Reducing the mid-frequency RT by half increased the STIPA value by 0.05 under the three configurations tested, agreeing with results presented by Mapp [5].

The subject based PB words speech test results are shown in figure 6 for three pairs of conditions (noise, no noise) for configuration: empty, with 10m² absorption and with 10m² with the subjects (native and non-native English speakers) facing away from the signal source.

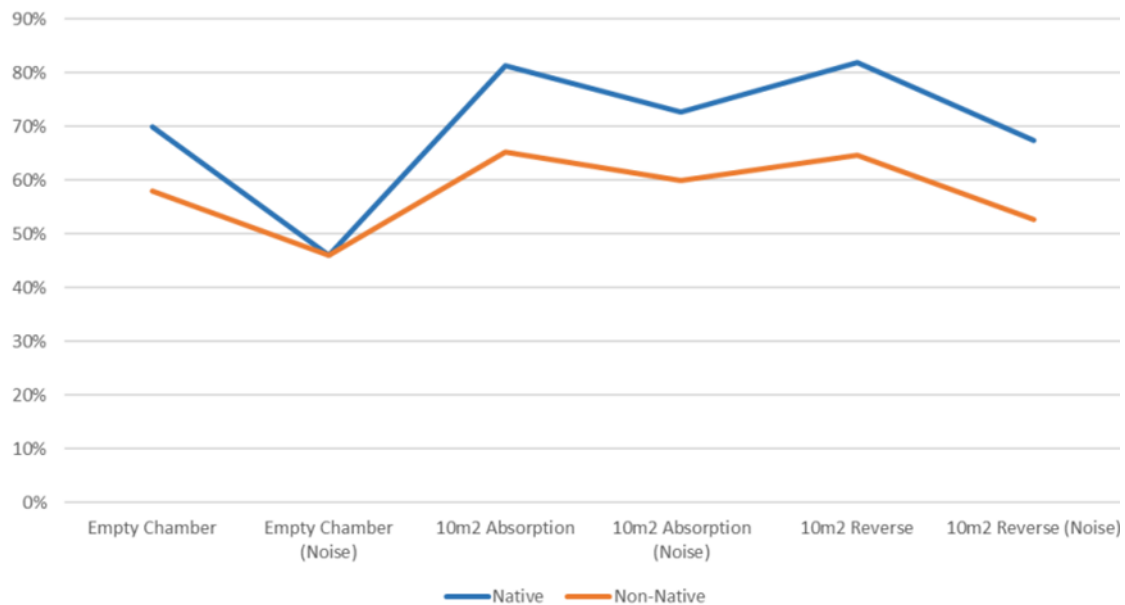


Figure 6 shows the averaged PB word scores for natives and non-natives under two room conditions.

From figure 6 it can be clearly seen that non-native listeners were not able to score as highly as natives in the speech intelligibility tests except in the absolute worst case condition when the scores were identical, indicating guessing of the test word, the corresponding STIPA value was 0.27. In all other conditions the natives consistently scored 12-16% higher. When considering native (no noise) vs non-native with noise there was a 20% increase in the PB word score, but this increased to 30% when the reversed listening condition was introduced. The reverse condition for the absorptive condition produced similar results +/- 1% for both natives and non-natives. However, adding pink noise to the absorptive condition created an overall reduction of 5% for natives and 7% for non-natives in the word score. This brought back the word score to the empty reverberant condition for native and non-natives. Thus, balancing out the 10m² of Class A absorptive treatment and then adding pink noise (SNR >15 to SNR = 8 dB).

The results can be compared to those given in ISO 9921 [2] which provides a suitable adjustment for non-native speakers, see table 2.

Table 2. ISO 9921 adjustments to the STI for different categories of non-native listeners

STI label range	Standard STI	Non-native category I experienced, daily second language use	Non-native category II intermediate experience and level of second language use	Non-native category III new learner, infrequent second language use
bad – poor	0,30	0,33	0,38	0,44
poor – fair	0,45	0,50	0,60	0,74
fair – good	0,60	0,68	0,86	impossible
good – excellent	0,75	0,86	impossible	impossible
For a non-native listener of category II and to achieve an intelligibility equivalent to an STI of 0,45, the transmission system needs to achieve a performance of 0,60. For intermediate values between the stated standard STI, interpolation should be used to estimate the adjusted STI.				
NOTE For details on STI label categories, refer to ISO 9921.				

A comparison of the STIPA measurements with the PB word scores is presented in figure 7.

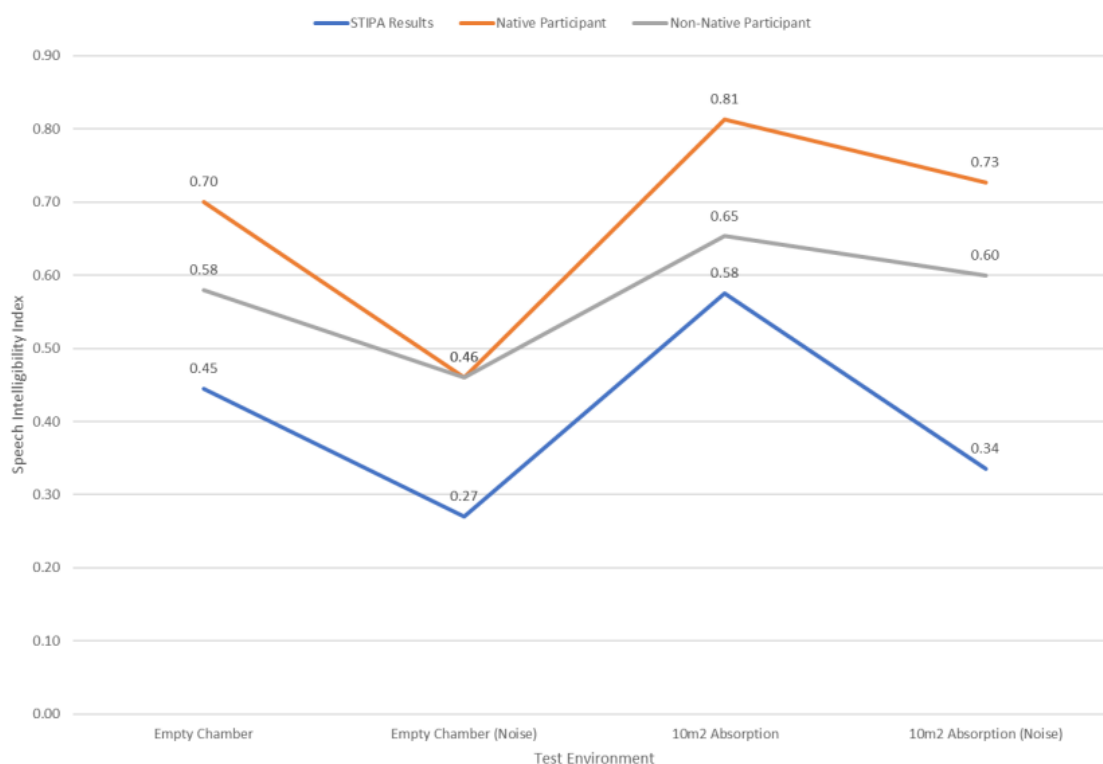


Figure 7. The objective and subjective (native and non-native listeners) results.

From table 2 and figure 7 it can be determined that the non-natives could be considered Category I – experience daily second language use [2]. From figure 7, the normal chamber condition (no additional absorption) produced a STIPA value of 0.45 without pink noise, the non-natives PB word score was 58% and the natives 70%, as would be expected. However, when noise was added STIPA dropped to 0.27 which produced a corresponding 46% PB score for both natives and non-natives, indicating that the subjects were guessing the test words, typically a STI of 0.3 equals a 50% PB score. When absorption was introduced, under the normal ambient condition STIPA increased to 0.58 resulting in an 81% PB score for native listeners and 65% for non-native.

4. CONCLUSIONS

An experiment was undertaken to determine the effect of changing the direction of the primary sound signal, such as you find on London Underground when typically, the loudspeakers are often positioned behind the passengers. A small sample of listeners, native and non-native, undertook an experiment under different noise and reverberation conditions. The results showed that both native and non-native listeners (Category I) were less able to determine the message when the loudspeaker was positioned behind, a reduction of 5% and 7% respectively, under a Signal to Noise ratio of 8 dB. There was only a marginal difference under excellent Signal to Noise ratio conditions.

5. ACKNOWLEDGEMENTS

This research was funded by London South Bank University and EPSRC UKAN+ (Grant number EP/V007866/1).

6. REFERENCES

1. BS EN IEC 60268-16:2020. Sound system equipment - Part 16: Objective rating of speech intelligibility by speech transmission index.
2. BS EN ISO 9921:2003. Ergonomics – Assessment of speech communication.
3. L. Morales, Validation and improvement of Speech Transmission Index method as a predictor of speech intelligibility, PhD Thesis, London South Bank University 2014
4. L. Morales, G. Leembruggen, S Dance, B. Shield, A revised speech spectrum for STI calculations, Applied Acoustics, 132, p33-42, 2018
<https://doi.org/10.1016/j.apacoust.2017.11.008>
5. P. Mapp 'Uncertainty in STI measurement and prediction' in Uncertainty in Acoustics, Ed R Peters, CRC Press, 2021