Acoustics 2008
Geelong, Victoria, Australia 24 to 26 November 2008

Acoustics and Sustainability:

How should acoustics adapt to meet future demands?

Effect of acoustic environment on the sensitivity of speech transmission index to source directivity

Ken Stewart and Densil Cabrera

Faculty of Architecture, Design and Planning, University of Sydney, NSW 2006, Australia

ABSTRACT

Source directivity has the potential to affect speech transmission index (STI) measurements. When the source is a model of a human talker, there is a question of how accurate that model needs to be. However, in many acoustic environments source directivity has little effect on STI because other factors dominate. One instance is soundfields in which the direct sound is dominant (e.g., in the quiet outdoors, in the nearfield, or in an anechoic room): directivity will have no effect on STI so long as the direct sound is strong enough. Another instance is soundfields in which the reverberant soundfield is dominant (e.g., in the far field in a room with moderate or more reverberation). This paper examines theoretical situations where source directivity has a substantial influence on STI because of the balance between direct (or early) and reverberant soundfields, as well as the role that background noise can have in increasing the importance of source directivity.

INTRODUCTION

Speech transmission index (STI) is an objective indicator of speech intelligibility derived from the measurement of the modulation transfer function of a system (International Electrotechnical Commission 2003). It may be used in assessing room acoustics, sound reinforcement systems, telecomunications systems, and other systems in which speech is conveyed to people. For a measurement often a loudspeaker would be used as the initial sound source for the test signal – and this loudspeaker should have similar directional characteristics to those of a human talker. However, in practice loudspeakers that closely match human speech directivity may not be available, so the question arises of how sensitive a measurement would be to deviations from ideal directivity. In this paper we consider the potential influence that sound source directivity has on STI in terms of source-receiver distance, room volume, room reverberation time, and background noise level.

Previous studies by Bozzoli et al. (2005a, 2005b) and Stewart and Cabrera (2007) have explored the issue of directivity of STI sound sources through physical measurement. By testing a variety of loudspeaker enclosures approximating the human form in various room acoustical contexts, Bozzoli et al. (2005a) found that directivity appeared to have little effect on STI measurements in normal rooms (classrooms), but did have a substantial effect in a car cabin. Bozzoli et al. (2005b) found that the measured average directivity of ten talkers was a good match for the directivity of a Bruel & Kjaer 4128C head and torso simulator. However, one issue with this is that a head and torso simulator is very expensive and somewhat fragile, and so is not necessarily a practical solution to STI measurement. Stewart and Cabrera (2007) examined the effect of mouth size on the directivity of a head and torso simulator, finding a substantial effect at some frequencies. When the effect of mouth size on STI measurements in real rooms (a lecture theatre and a meeting room) was tested, there was little effect, with only a modest variation in STI in one of the positions tested.

These measurement-based studies are limited by the effort involved in making physical measurements, and so only offer a glimpse of the range of possible scenarios. Hence the present paper takes a theoretical appoach to the question.

MODEL OF STI

According to Houtgast et al. (1980), the modulation transfer function of a room acoustical system, $m(F)$, can be estimated from equation 1, based on principles from statistical room acoustics. The modulation transfer function is formed by a set of moduation reduction coefficients (i.e., values that enumerate the extent to which the modulation depth of a modulated signal is reduced by the acoustic system) for a range of modulation frequencies (F), which for STI measurements span 0.63 Hz to 12.5 Hz in 1/3-octave intervals. In STI measurements, this modulation transfer function is calculated for each of seven carrier frequency octave bands, from 125 Hz to 8 kHz.

$$m(F) = \frac{\sqrt{A^2 + B^2}}{C}$$

$$A = \frac{Q_t Q_l}{r^2} + \frac{1}{r_c^2\left[1 + \left(2\pi FT/_{13.8}\right)^2\right]}$$

$$B = \frac{2\pi FT}{13.8 r_c^2\left[1 + \left(2\pi FT/_{13.8}\right)^2\right]}$$

$$C = \frac{Q_t Q_l}{r^2} + \frac{1}{r_c^2} + Q_t \times 10^{\left(L_N - L_{t,1m}\right)/10} \tag{1}$$

Here Q_t is the directivity factor of the talker, and Q_l is the directivity factor of the listener. In emulating STI measurements Q_t is the directivity of the measurement louspeaker and Q_l=1 (an omidirectional measurement microphone is used). For a sound source, directivity factor is the ratio of sound intensity radiated in the reference direction to the average sound intensity radiated in all directions at a given distance from the source. Usually the on-axis direction (i.e. the normal to the face of the sound source) is taken as the reference direction. Alternatively, directivity index, *DI*, is directivity factor expressed in decibels, i.e.

$$DI = 10\log Q \tag{2}$$

Returning to equation 1, the source-receiver distance is *r*. The critical distance (the distance from a sustained omnidirectional source for which the direct and reverberant fields contribute equally to intensity) is denoted r_c. *T* is the reverberation time of the room at the carrier frequency under consideration. L_N is the background noise level, and $L_{t,1m}$ is the long term equivalent sound pressure level of the speaker at a distance of 1 m. This level is standardised in IEC60268-16(2003) as total of 60 dB(A) summed across the seven (for male speech) or six (for female speech) octave bands.

Each of these modulation reduction coefficients is converted to an 'apparent signal-to-noise ratio', $(S/N)_{app,F,oct}$, stated in decibels (equation 3). The subscript 'oct' refers to the fact that this calculation would normally be performed for each octave band carrier frequency (as well as each modulation frequency, *F*). This value is influenced by both the actual noise level and the temporal smearing caused by reverberation. It is determined in the following manner:

$$(S/N)_{app,F,oct} = 10\log\frac{m(F)_{oct}}{1 - m(F)_{oct}} \tag{3}$$

These values are then clipped, so that values greater than 15 dB are reduced to 15 dB, and values less than -15 dB are increased to -15 dB.

The mean apparent signal-to-noise ratio for each carrier octave band is calculated according to equation 4:

$$(S/N)_{app,oct} = \frac{\sum (S/N)_{app,F,oct}}{n_F} \tag{4}$$

Here, n_F is the number of values (modulation frequencies) represented in a carrier octave band, which is equal to 14 for STI measurements (i.e., the 14 frequencies from 0.63 Hz to 12.5 Hz).

The apparent signal-to-noise ratios (one for each of the seven octave band carrier frequencies in STI measurements), are converted to sound transmission indices according to equation 5.

$$STI_{oct} = \frac{(S/N)_{app,oct} + 15}{30} \tag{5}$$

Although a full calculation of STI would then involve a weighted combination of seven STI_{oct} values, taking into account interactions between the bands due to masking, for the modeling in this paper we omit these final steps. This allows us to consider just one reverberation time and background noise level for a given output value (instead of reverberation time and noise varying between octave bands), and so provides a straightforward way of modelling the effect of directivity on STI. In the following discussion of our modeling, we use the term 'STI' to refer to what would be an STI_{oct} value in a full calculation.

STI has a value between 0 and 1, where 1 corresponds to good transmission. Depending on the context, a value of 0.45 (the lower limit of 'fair') and higher may represent acceptable intelligibility.

CALCULATED STI AND SENSITIVITY TO DIRECTIVITY FACTOR

In order to characterise the effect of source directivity on STI we have obtained values from the model described in the previous section for directivity factors (Q_t) ranging between 0.125 and 8 using logarithmic steps (or directivity indices between -9 and 9 dB using linear steps). Of course it would be most unusual to intentionally employ a sound source with a directivity factor of less than 1 (meaning that the source is directional, but its energy is radiated predominantly in a direction other than the reference direction – for example when a person speaks with their back facing the listener). On the other hand, there are situations where some flexibility in the talker's orientation is desirable, so it may be useful to design against potential problems arising from directivity indices of less than 1. In a later section of the paper we consider the range of directivities (as a function of frequency) for human talkers and two electroacoustic sound sources designed for speech emulation.

We consider two cases: (i) the effect on STI of changing source directivity whilst maintaining constant on-axis free field sound pressure level at 1 m; and (ii) the effect on STI of changing source directivity whilst maintaining constant source power. The first case is relevant to considering the extent to which the directivities of various potential sound sources for STI measurement would influence the result. The second case is relevent to considering the effect of source orientation on STI.

The standard deviation of STI for $0.125 \leq Q_t \leq 8$ was taken as the indicator of sensitivity of STI to directivity. An alternative approach might have been to take the difference between maximum and minimum (which would have yielded larger values), but we prefer standard deviation because it is derived from the full set of calculated values rather than the two extremes.

For the range of directivity indices indicated above, we then determined STI considering the following parameters: source-receiver distance (*r*), reverberation time (*T*), room volume (which together with reverberation time affects critical distance r_c), and background noise level (L_N). Three room volumes were considered – 130 m^3, 1300 m^3 and 13000 m^3. Values of *r* were scaled by the cube root of room volume (0.05 m – 6.4 m in the 130 m^3 room, 0.11 m – 13.79 m in the

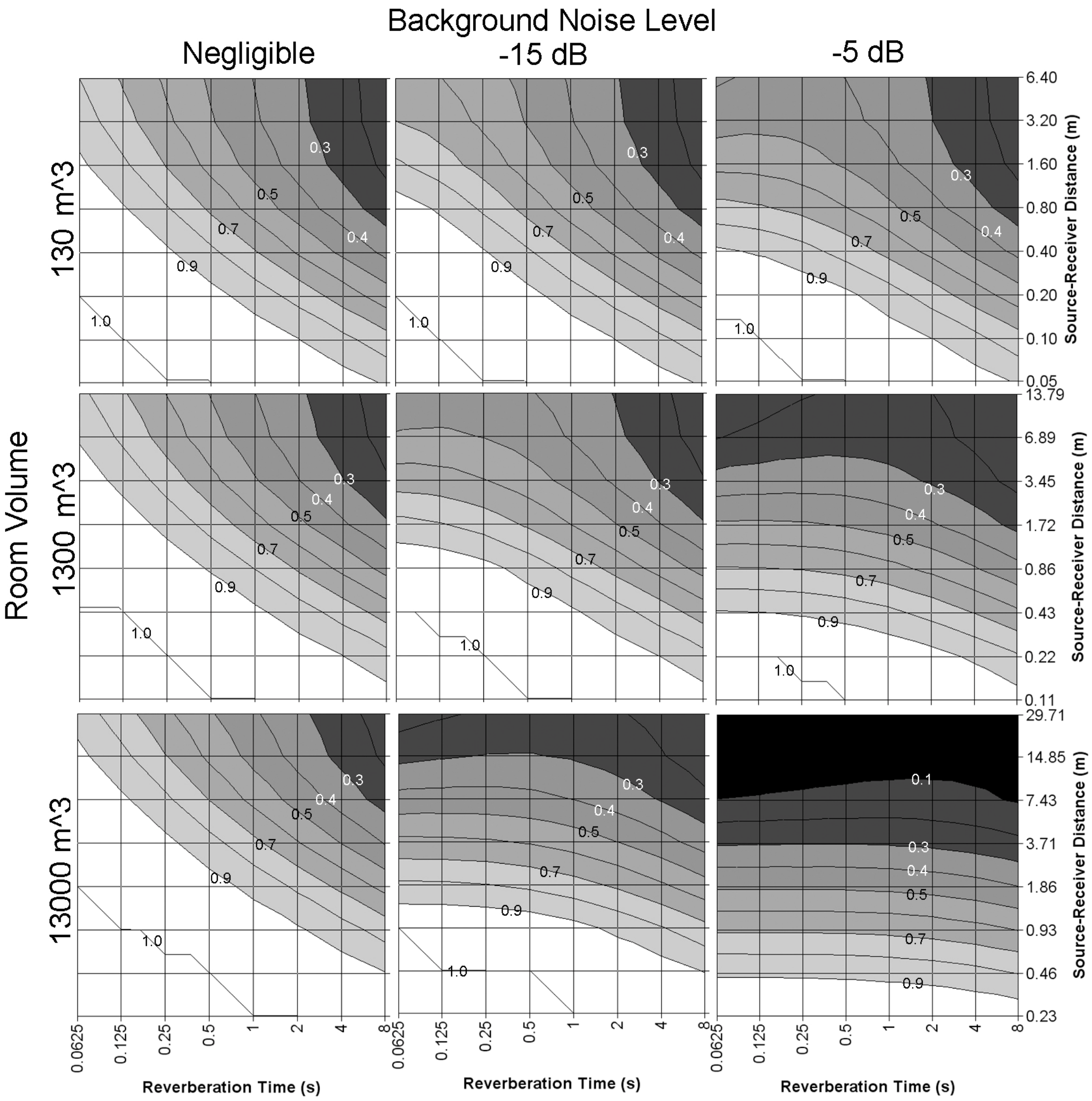

Figure 1. Mean STI for a range of source directivity factors (Q_t) from 0.125 to 8. Results are shown as a function of reverberation time, source-receiver distance, room volume and background noise relative to anechoic on-axis speech at 1 m. While values are for the case of constant on-axis pressure, almost identical values are found for the case of constant source power.

1300 m^3 room, and 0.23 m – 29.71 m in the 13000 m^3 room). Values reverberation time ranged between 0.0625 s and 8 s (not scaled by room volume). Three background noise levels were considered: negligible noise, -15 dB relative to anechoic speech at 1 m (which could be interpreted as being equivalent to 45 dBA), and -5 dB relative to anechoic speech at 1 m (i.e. equivalent to 55 dBA).

Mean values of STI for the range of directivity factors are shown in Figure 1 in terms of source-receiver distance, reverberation time, room volume and background noise level. Almost identical mean values are found for the two cases (constant on-axis free field pressure, and constant power), so only the first of these cases is presented here. Figures 2 and 3 show standard deviations of STI due to variation of source directivity factor, for each of the two cases. The values of Figure 1 are mainly provided to help interpret Figures 2 and 3.

Effect of directivity in low background noise

Results for the negligible background noise condition are shown in column 1 of Figures 1-3. As would be expected, a long reverberation time and large source-receiver distance yields low STI values (Figure 1, column 1). There is no difference in the effect of directivity on STI between the two source conditions (constant pressure in Figure 2 versus constant power in Figure 3). Those sub-charts show that sensitivity to directivity reaches a maximum at relatively short source-receiver distances, and that this peak sensitivity increases with reverberation time. The source-receiver distance for peak sensitivity decreases with increasing reverberation time, as might be expected from the reduction in critical distance. However peak sensitivity is not at the critical distance, but instead is around 60% of the critical distance (however, this should not be taken as a general principle because it depends on the range of directivity factors evaluated). Scaling the source-receiver distance by the cube root of room volume results in a similar pattern of sensitivities for the three room volumes evaluated.

Effect of background noise on directivity sensitivity

Introducing background noise makes STI less dependent on reverberation time, especially in the larger volume rooms (Figure 1, columns 2 and 3). In the case of constant source pressure (Figure 2) background noise introduces a second peak area of sensitivity to source directivity, at large source-receiver distances. This area of peak sensitivity moves from short to longer reverberation times as the background noise

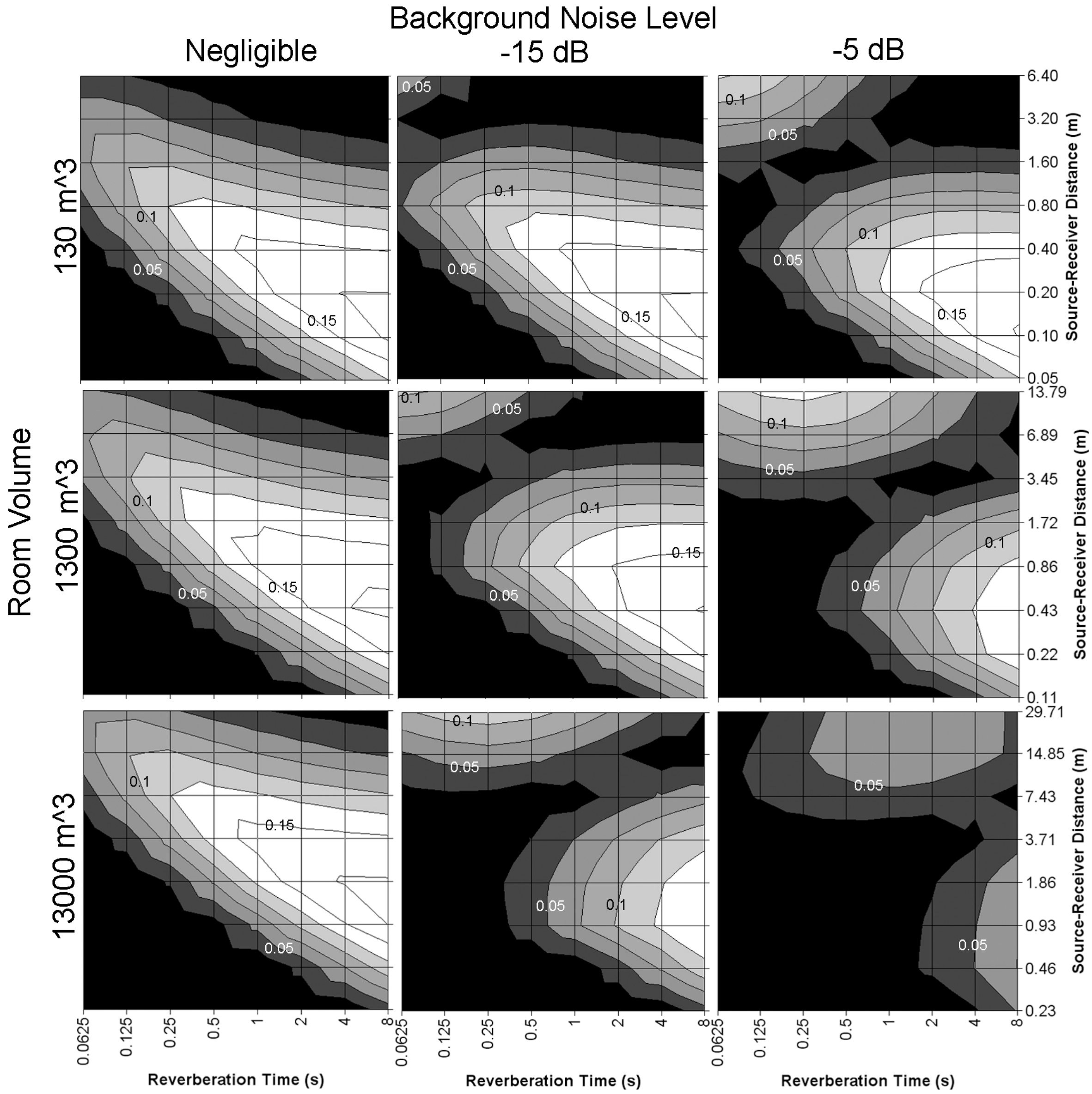

Figure 2. Standard deviation of STI for a range of source directivity factors (Q_t) from 0.125 to 8, for the case of constant on-axis sound pressure. Results are shown as a function of reverberation time, source-receiver distance, room volume and background noise relative to anechoic on-axis speech at 1 m.

level increases, and as the room volume increases. Hence this appears as an interaction effect between background noise and reverberation, rather than an effect of noise alone. The reason for this second peak is that in varying the directivity factor, the model is maintaining the on-axis anechoic sound pressure level – hence the power of the source increases as the directivity decreases. In the second peak area, it is this consequent change in the power of the source, rather than directly its directivity, that is contributing to higher STI values for a given room acoustical and background noise condition. In the large room conditions, the second peak area occurs in areas of relatively low STI (consider Figure 2 in relation to Figure 1), in some cases much lower than a practically useful value for speech communication. In the small room STI values are reasonably good in the second peak area, and so it may become relevant to understanding smaller spaces.

In the case of maintaining constant source power, there is no second peak area of sensitivity, and the introduction of background noise increases the sensitivity of STI to directivity at short reverberation times (Figure 3).

Floor and ceiling effects are important in influencing the dark areas (low standard deviations) of Figures 2 and 3. In conditions of very high mean STI, or very low mean STI there is little scope for directivity to influence STI. However, peak sensitivity is not generally at a mean STI of 0.5, but at 0.7--0.8 in rooms with negligible noise.

An important consideration here is that the range of directivity factors that were evaluated (Q_t of 0.125 to 8) influences both the mean and the standard deviation, and for some purposes, evaluating directivity factors of less than 1 is much less relevant than higher directivity factors – especially in the first case (constant on-axis pressure), which might be interpreted as being relevant to the selection of loudspeakers for STI evaluation. If larger Q_t values are evaluated, then the peak sensitivity is shifted to greater distances, and potentially beyond the critical distance. This point is considered further towards the end of the paper.

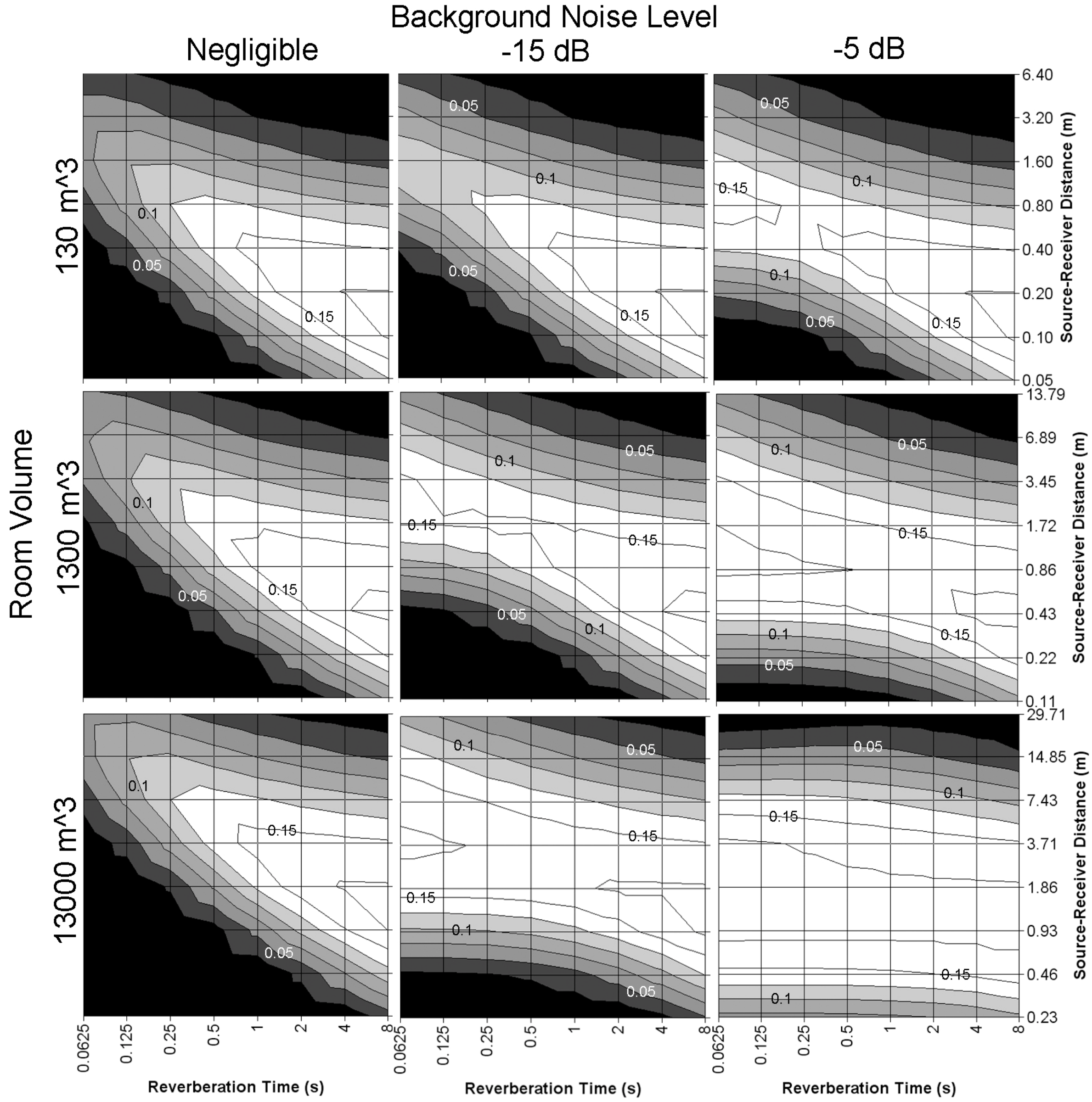

Figure 3. Standard deviation of STI for a range of source directivity factors (Q_t) from 0.125 to 8, for the case of constant power. Results are shown as a function of reverberation time, source-receiver distance, room volume and background noise relative to anechoic speech at 1 m when $Q_t = 1$.

DIRECTIVITY INDICES OF SPEECH SOURCES

On-axis

The interpretation of the results of this model requires some knowledge of the directivity indices of speech sources. In this section we consider the on-axis directivities of people, as well as two electroacoustic sources that might be used for STI measurement.

The most extensive measurements available for human speech directivity come from a study by Chu and Warnock (2002). For 40 subjects, they measured long term average sound pressure levels of conversational speech in 105 directions around the talker (with 9 elevation angles and 13 azimuth angles). They also performed the same measurements with a Bruel and Kjaer type 4128C head and torso simulator. To use their results in this study, we have collapsed their measurements into directivity factor values. Results for on-axis directivities are shown in Figure 5.

For this study we also measured the directivity of a loudspeaker designed for a type of STI measurement – the NTI Talkbox. This loudspeaker is part of a kit for the direct measurement of STIPA (a simplified version of STI intended particularly for measuring public address systems).

In the context of a public address sytem, the measurement loudspeaker would be put at a person's (e.g., announcer's) position with the same relationship to the system's microphone as the real person – thereby factoring in the acoustic conditions of an announcement booth for example. Nevertheless, this kit is also a convenient tool for measuring STIPA in purely acoustical contexts, and indeed may provide more robust measurements than RASTI (Room Acoustics Speech Transmission Index) in environments where reverberation time varies significantly across the 125 Hz – 8 kHz frequency range. This kit is smaller, cheaper and more rugged than a Bruel & Kjaer head and torso simulator with an artificial mouth.

Figure 4. NTI Talkbox directivity measured using Bruel & Kjaer Electro-acoustic Pulse System Type 7907 with turntable in Anechoic Chamber

Measurements were made in an anechoic room at a distance of 1.5 m. The loudspeaker was rotated using a turntable, with measurements made every 10 degrees. Microphone elevation angles were at 20 degree intervals, ranging from -40 degrees from the horizontal to 80 degrees (90 degrees was also measured). Measurements were made using the Bruel & Kjaer steady state response Electro-acoustic Pulse System Type 7907. A photograph from the measurement is shown in Figure 4, and results for on-axis directivity are shown in Figure 5.

Following Beranek (1986), the directivity of simple circular piston sources can be predicted based on the product of wave number and their radius (which is equivalent to the ratio of circumference to wavelength). Values for an 80 mm diameter piston on the end of a long cylinder are shown by way of comparison to the measured values of the NTI Talkbox on Figure 5 (the Talkbox has a loudspeaker diaphragm diameter of about 80 mm). This shows greater directivity for the NTI Talkbox over much of the frequency range, presumably because of the physical effect of the rectangular loudspeaker enclosure. In the high frequency range it is usual for the loudspeaker cone to mainly radiate from the centre, reducing its effective diameter, which would counteract the added directivity due to the box – and this is seen in the high frequency results.

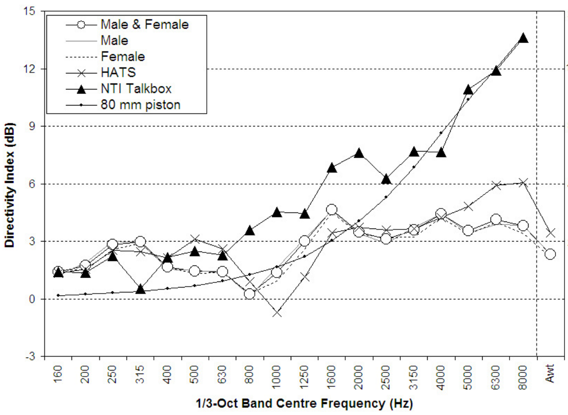

Figure 5. On-axis directivities of human talkers and a Bruel & Kjaer 4128C head and torso simulator, derived from Chu and Warnock's (2002) data; and the measured on-axis directivity of an NTI Talkbox, compared with the expected value for an 80 mm circular piston on the end of a long cylinder.

Based on the values in Figure 5, we can expect directivity index values between about 0 dB and 5 dB for real people, or -1 dB and 6 dB for a 4128C head and torso simulator. When a more conventional loudspeaker is used, directivity is likely to be greater in the high frequency range, with the NTI Talkbox attaining a directivity index greater that 12 dB in the very high frequency range. In the vicinity of 2 kHz (the part of the spectrum to which speech intelligibility is particularly sensitive) the NTI Talkbox had a directivity index approximately 3 dB greater (i.e. double directivity factor) than the humans and 4128C.

Effect of source rotation on directivity index

Directivity index can be calculated for any reference angle, and in this section we consider the directivities that occur as a source rotates around the horizontal plane. Figure 6 shows the directivity indices for Chu and Warnock's human data and the head and torso simulator, with angle of rotation on the vertical axis. This shows that, in the high frequency range, there can be substantially negative directivity indices, which our modelling of STI standard deviation suggests can be important in some circumstances for speech intelligibility.

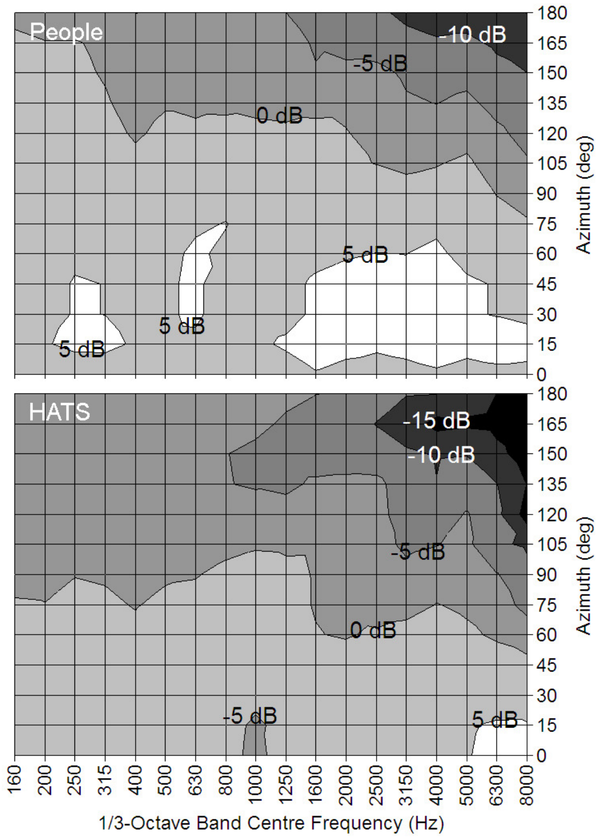

Figure 6. Measured directivity index of people speaking, and of a head and torso simulator (HATS) as a function of frequency and angle of rotation (based on Chu and Warnock, 2002).

Figure 7 shows directivity index for rotations of the NTI Talkbox. As would be expected, there are much more substantial effects on directivity index as the loudspeaker rotates, especially in the high frequency range.

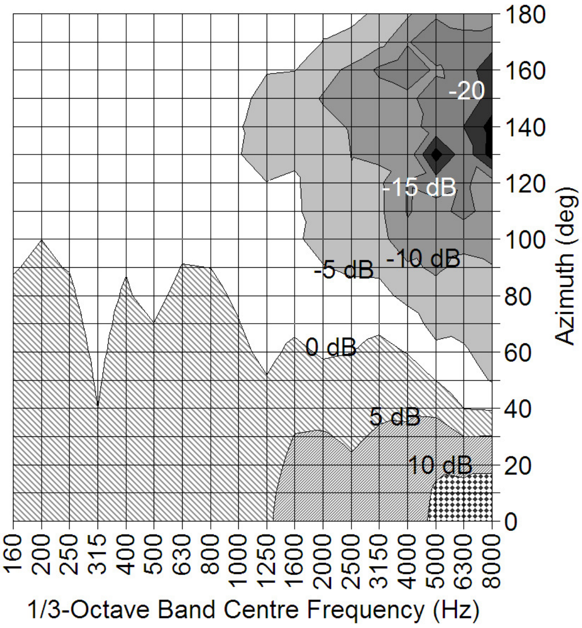

Figure 7. Measured directivity index of an NTI Talkbox as a function of frequency and angle of rotation.

DISCUSSION

To interpret these results in terms of the standard deviations in the first part of the paper, we should consider the range of values used to generate the standard deviations (-9 dB to 9 dB). Variations in STI of less than 0.05 are likely to be too small to be of consequence. However, the largest standard deviations are of the order of 0.2, which, of course represent a wider range of STI values (the corresponding range between maximum and minimum is 0.57 in some cases). However, the range of directivity index values seen as loudspeakers are rotated can extend substantially beyond the -9 dB to 9 dB range, and so can have a large effect on octave band STI under the right combination of acoustical conditions. Rotating a loudspeaker corresponds to the constant power case.

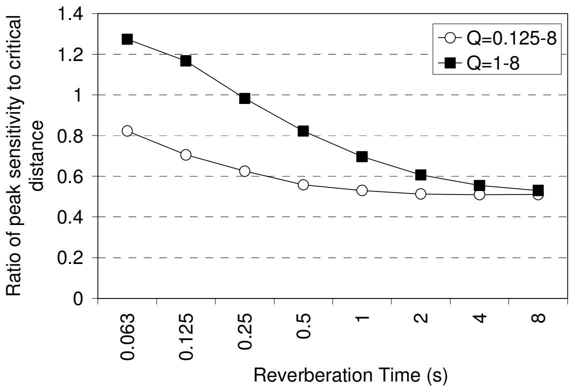

Figure 8. The distance of peak sensitivity to source directivity divided by the critical distance as a function of reverberation time. Two ranges of directivities are evaluated: Q_t = 0.125 – 8; and Q_t = 1-8.

On the other hand, in the case of constant on-axis pressure, it would be unlikely that negative directivity index values would be used in realistic applications. Figure 8 shows the relationship between the distance that has peak sensitivity to directivity to the critical distance of the room for the range of directivity factors evaluated for Figures 1-3, and also for a range of directivity factors between 1 and 8 (noise-free conditions). Note that these relationships are independent of room volume, and the same for constant pressure and constant power cases.

CONCLUSIONS

Previous studies have found that source directivity often has little effect on STI measurements, but can have a substantial effects in some circumstances. This paper has looked at the acoustical conditions for which source directivity has maximum influence on speech transmission index. Maximum sensitivity is related to the critical distance and reverberation time in noise-free conditions. A second area of sensitivity is seen when noise is introduced (assuming on-axis pressure remains constant while directivity changes). The applications for this work include prioritising issues in the design and selection of artificial speakers for STI measurement, as well as understanding speech intelligibility over short source-receiver distances (for example, in meeting rooms or restaurants).

REFERENCES

Beranek, L.L., 1986, *Acoustics*, American Institute of Physics.

Bozzoli, F., Bilzi, P. and Farina, A. 2005a, "Influence of artificial mouth's directivity in determining speech transmission index," *119th Audio Eng. Soc. Conv.*, New York, USA

Bozzoli, F., Viktrorovitch, M. and Farina, A. 2005b "Balloons of directivity of real and artificial mouth used in determining speech transmission index," *118th Audio Eng. Soc. Conv.*, Barcelona, Spain

Chu, W.T. and Warnock, A.C.C. 2002, *Detailed Directivity of Sound Fields around Human Talkers*, National Research Council Canada IRC-RR-104

Houtgast, T., Steeneken, H. and Plomp, R. 1980, "Predicting Speech Intelligibility in Rooms from the Modulation Transfer Function," *Acta Acustica united with Acustica* 46, 60-81

International Electrotechnical Commission, 2003, *Sound system equipment – Part 16: Objective rating of speech intelligibility by speech transmission index*, IEC 60268-16

Stewart, K. and Cabrera, D. 2007, "Effect of artificial mouth size on speech transmission index," *14th International Congress on Sound and Vibration*, Cairns, Australia