Speech Intelligibility Evaluation for Mosques

Mokhtar Harun, Tharek Abdul Rahman, Md. Najib Ibrahim, Ahmad Khan Said*

Faculty of Electrical Engineering
**Faculty of Built Environment*
Universiti Teknologi Malaysia
81310 UTM Skudai Johor.
Tel: 07-5576160 ext 5158
E-mail: mokhtar@suria.fke.utm.my

Abstract- **Unlike mosques, few of enclosed room functionality concentrates only on speech. Even though large contemporary mosques are built into rooms such as multipurpose hall, lecture hall, dining hall and library, main prayer areas are still the most vital room in mosques. With least tolerance to appreciable level of background noise, with volume easily exceeded 10,000 cubic-meter, hard and sound reflecting surfaces are everywhere, and with the installation of dome or domes, mosques should be critically design to manifest, besides its aesthetics and thermal comfort values, an excellent acoustics. It is intended for the good of concentration and serenity of worshippers who come into mosques to dwell temporarily for any religious activities. Because of its sole functionality for speech, speech intelligibility evaluation in mosques requires a set of selective approaches compared to that in multi functional enclosed rooms. This paper discusses architectural acoustics of mosques and proposes their speech intelligibility evaluation techniques.**

I. INTRODUCTION

Mosque, apart from its various and profound roles for Muslim community, is a "listening room" which desires the highest quality listening conditions. Speech intelligibility in mosques will manifest to be good or poor as a result of the following two combinations: the shape and form of the room that contains it and a fine-tuned audio system [1].

Unlike pure tone, speech and background noises are derived from many different frequencies occurring simultaneously and are rich in harmonics. Background noise can be originated from the sound of nearby traffics, and air-conditioning and mechanical ventilating system.

Speech may be coming from the *imam* preaching from the *mihrab* or speech sound of localized group's *usrah* that is taking place scattered in all over the floor in the main prayer hall. Background noise and speech include a wide range of frequencies, normally from 20 Hz to 5000 Hz, will be dispersed throughout the space in mosque, until it attenuates and decays through time.

The direct sound from the talker that reaches the listener in less than 35 ms is crucial to speech intelligibility. With contemporary mosques built with levels to the side and at the back of the main prayer area, there too, should receive adequate amount of direct sound for excellent intelligibility.

With the assumption that the installation of audio system is electronically flawless, distance of loudspeakers to the listeners should be properly adjusted. They are to cater group of listeners, with the precaution that the main portion of sound radiation of loudspeakers should not focus on the reflective wall or untreated floor [2].

The diffused and uniform dispersion of sound and its time of decay that affect speech intelligibility in mosques are the fundamental discussion of this paper. These two characteristics of sound affect the intelligibility of speech in many ways.

On the other hand, human ears are not equally sensitive to sound occurring at all frequencies, but rather are most sensitive to sounds between 1000 Hz and 5000 Hz. Besides, the percent contribution of each of the frequency bands that contribute to the articulation is 91.5% between 200 Hz to 4000 Hz [1]. Therefore, the frequency range for this paper will be limited from 125 Hz to 5000 Hz, with the interval frequencies concerned follow either one-octave or 1/3-octave center frequencies.

II. ACOUSTICS IN MOSQUES

a. Dispersion of Sound

Sound waves behave quite differently with respect to their frequencies. For sound waves below 300 Hz - 400 Hz, these very long wavelengths of can diffract and bend around most objects in the room. Therefore at low frequencies, sound wave is "non-directional". As it move up the frequency spectrum, sound waves move freely less affected by the object in a room [3].

At the middle and high frequency sound waves are more directional and tend to act in a specific geometric manner. They move in path like a ray of light.

As the sound waves follow the path like a ray of light, there will be angle of incidence and angle of reflection, at point where the wave strikes the material it confronts and be reflected back throughout the room. This specular reflection, where some original energy of wave is being reflected back will continue to move around the spaces until the wave is attenuated to a point where it will no longer be audible. The acoustics of small and large listening room will suffer more if there is symmetrical, parallel, rough and hard surfaces in the room

b. Decay time of sound

Mosque is categorized as large listening room where its volume easily exceeding 10,000 m^3. Unlike small listening room, acoustics quality in mosque is less affected by resonance, which in the form of standing waves but it is more by reverberation [4].

Even though reverberation is not the only parameter in evaluating mosque's acoustics, it is still the most important. The amount of time for sound reflections to decay 60dB from its initial value or reverberation is really significant or useful to add fullness or warmth to music by allowing the sound in the room to blend [3][4].

Longer decay of sound will let the music "linger" in the air before it dies. For mosque in which speech is the main acoustical function, longer reverberation causes masking of critical elements of speech that pollutes the intelligibility.

By using Sabine's formula

$$RT60 = \frac{0.16V}{A} \qquad (1)$$

Where V is volume in meter and A stands for total effective absorption of sound energy, Sabine $-m^2$, in a room.

As the volume increases, reverberation time will also increase. Surprisingly, in mosque, absorptive materials are hard to be installed. This is due to the fact that the structural strength and aesthetics values require stone, and hard reflective material to be furnished. It is not common to have pillars made of soft rockwool.

In addition, decoration, in forms of Quranic versus and cornices are more practical to be carved of hard material. Decorations on hard surfaces are easy to maintain and will stand the test of weather and time.

Moreover, with current large mosques are fully air-conditioned, the application of permanent carpet on small prayer area is questionable. This is due to the fact that the carpet will retain moisture, and this over time will produce uncomfortable odor.

III. SPEECH

Good acoustics for speech are generally incompatible with music. Reinforcing music has only one major criterion - it must be loud and clear, which involves transducers that translate sound into a lot of electrical power [1]

For auditoria, if the speech is intelligible and the background noise is not intrusive, audience's dissatisfaction is unlikely [3]. Optimum reverberation time, good sight line, elimination of echoes and keeping audience close to the stage are standard requirements of any auditorium for good speech intelligibility.

Speech, in any language, generally consists of vowels and consonants. The former is commonly louder and of longer duration and its delay is in 90 ms that the latter is shorter in nature, which lasts about 20 ms.

As in Figure 1, greater energy in vowel has produced sound masking upon consonants. Human speakers, loudspeakers and most other sound sources are virtually omni-directional at low frequency and become progressively directional at high frequencies. This explains why sound speech at high frequencies gather near in front of the speaker. Therefore, the directivity of the high frequency component of sound is the most affected.

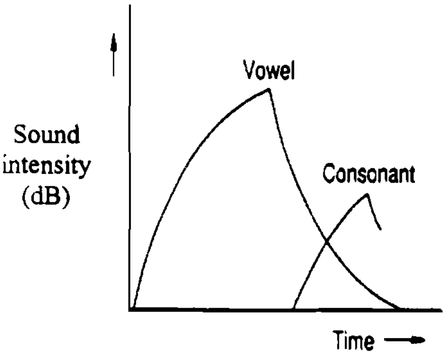

Figure 1 The sound intensity versus time history in a room for a vowel followed by a consonant (after Barron, 1998)

Speech power is a function of frequency and its power distribution is not linear over the frequency spectrum. As in Figure 2, it can be seen that providing a smooth frequency response, through the very critical area from 250 Hz to 5000 Hz, is much more important than extending the frequency response of a transducer. Speech frequency response works well within 200 Hz to 4000 Hz [1]. It is obvious that small radio with small speakers works just as well when compared to a high fidelity music system with wide frequency response.

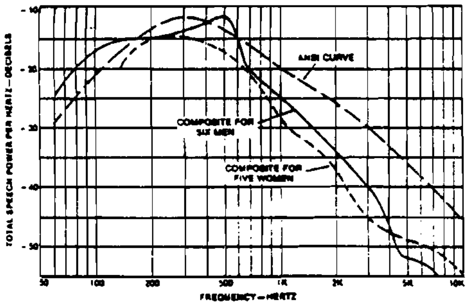

Figure 2 Relative speech power as a function of frequency for men and women (after Davis *et.al*, 1997)

Also, from Figure 2,it was found that from the composite for six men voice levels, the low frequency responses remain almost the same. The voice power spectrum has been uniformly increased until 500 Hz and the power start to decrease from middle to high frequency. It can be said that the louder voice spectrum is suitable only for projection of sound but not for speech intelligibility. It is obvious from Figure 3 that highest articulation contribution which is 11%, is at 2000 Hz. Little intelligibility information is being contributed by speech power spectrum at low frequency, but by adding the percent of contribution to the articulation index, 91.5% of the articulation is contributed by frequency bands between 200 Hz to 4000 Hz.

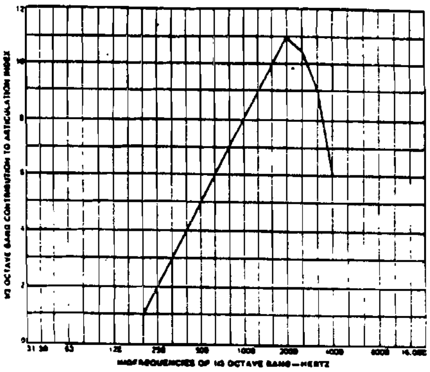

Figure 3 Variation of an articulation index contribution with speech components in 1/3-octaves (after Davis *et.al*, 1997)

IV. SPEECH INTELLIGIBILITY

Barron has showed, through the subjective survey of British concert halls, that listeners were responding consistently on three attributes: intelligibility, intimacy and reverberant. The crucial result was that only intelligibility was rated significant for judgement of the acoustics overall. In another point listeners were able to assess acoustics mainly on intimacy and reverberant but were indifferent to the degree of these attributes [3].

Loudness that is the net reflection of sound, even though has been an acoustic concern, is not felt to be significant in the speech intelligibility requirement for as long as it is loud enough to fill the room relative to background noise [3]. Ironically, for a room with considerable long reverberation time, the reverberant level in that room will increase. This reverberant level will also constitute itself as if it is the background noise as seen in Figure 4. The distance R, which is about 5 meter for 10,000 cubic-meter with RT60 = 1.5s is to disclose that what one is listening to in a large room mostly is the reverberant sound. R is given by [5],

$$R = 0.1[V/\pi T]^{1/2} \qquad (3)$$

Where

R = distance at which reverberant and direct Sound level are equal, m
V = room's volume, m^3
T = reverberation time, s

Thus, speech intelligibility is mainly a function of sound level, distance from the source and the characteristics of the room that contains it, since subjective loudness is a function of these three aspects.

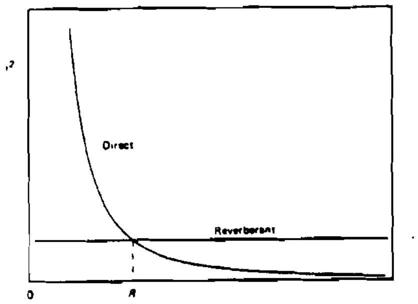

Figure 4 The distance R at which the direct sound and reverberant level are equal (after Hall, 1993)

Sound level in forms of early energy traction is necessary, requiring a high early energy and low late energy. High early energy such as direct sound, and according to the inverse square raw, decreases 6 dB every doubling of distance.

Closer distance of the listener from the source ensures that the listener receives as much early energy as possible.

The effect of late reflections or late energy components is reduced with the high average absorption of the room. Characteristics of the room such as sloped seats, balcony, curved surfaces tend to create the effect of shadowing and focussing. These, in turn, will affect the reverberation time and uniform diffusion of sound, and thus resulting variations in speech intelligibility. However the behavior of the late reflected sound in a room is more consistent and less influenced by the orientation and position of acoustically reflecting surfaces [3]. It is the uniformity of sound distribution that is affected by these characteristics.

V. SPEECH INTELLIGIBILITY EVALUATOR FOR MOSQUES

Several techniques for speech intelligibility are now available for practical uses. There are techniques such as articulation Index (AI), Rhyme Test, Sound Transmission Index (STI) and Rapid Articulation STI (RASTI).

Direct percentage Alcons (percentage loss of consonants) testing, although based on concepts dating back to the 1960s and early 1970s, is in fact a relatively new, measurement technique. The method measures the ratio of the direct to reverberant sound components received from a sound system or test loudspeaker at a typical listening position. The measurement is carried out at the frequency 2000 Hz only. However, measurements at other frequencies are often carried out to give a more detailed picture of a particular system.

This paper proposes among other, the Articulation Loss of Consonants (Alcons) as speech evaluator for mosques due to the following reasons.

First, Alcons, covers various manipulative intelligibility indicators such as reverberant sound level, direct sound level, ambient noise level and reverberation time. Unlike AI, Alcons is not only able to predict the effect of background noise on speech intelligibility but also the behavior of speech sounds transmission itself [3][7].

Second, Alcons employs only 2000 Hz frequency band as its intelligibility indicator, at which it contributes 11% or the highest percent of articulation [1].

Third, by employing Alcons method, speech intelligibility evaluation can be assessed both by objective and subjective means. Unlike Rhyme Test, which is a solely subjective intelligibility evaluator, requires both speakers and listeners to be native English speakers.

Fourth, as proved by Puetz (1978), Alcons is outstanding to be used in listening room with volume of more than 10,000 m^3, in which reverberation time is quite higher so as to normally, be more than 2s [6].

Fifth, the interpretation of intelligibility results of Alcons reflects much on human hearing capabilities, at par with the functionality of mosques, which are mainly

on speech. Although RASTI and STI are the latest speech intelligibility evaluator, their input and output are a result of complex computer program, and thus their parameters are less likely manipulative for any other analysis.
Furthermore, one of the most useful features of the Alcons method is that one can correlate measurement with prediction [7]. Alcons is still effectively the only method enabling of predicting the potential intelligibility of a sound system before it is installed.

Follows are the established simple formula that was introduced by Puetz in 1978.

$$\% \text{ Alcons} = \frac{200D^2T^2(n+1)+K}{QV_m}$$

Alternatively:

$$\% \text{ Alcons} = 100\ (10^{-2([A+B)-ABC])} + 0.015$$

where

$$A = -0.32 \log\left(\frac{L_R + L_N}{10L_D + L_R + L_N}\right) \text{for } A \geq 1,\ A = 1$$

$$B = -0.32 \log\left(\frac{L_N}{10L_R + L_N}\right) \text{for } B \geq 1,\ B = 1$$

$$C = -0.5 \log\left(\frac{R_{T60}}{12}\right)$$

Where :

D = listener distance to source (m)

T = Reverberation time (s)

n+1 = number of like sources (group of loudspeakers) contributing to the reverberant field.

Q = loudspeaker directivity (axial factor)

V = volume of the space, m^3

M = critical distance modifier $= \frac{1+\alpha}{1-c}$

α = average absorption coefficient of the room

c = average absorption coefficient of surface covered by the loudspeakers

K = listener correction factor, e.g. 1.5 % for good listener

L_R = reverberant sound level (dB power ratio)

L_D = direct sound level (dB power ratio)

L_N = ambient noise level (dB power ratio)

R_{T60} = reverberation time (s)

The concept works more as a series of bands:

> 15%	not acceptable except for very simple or well known messages
15-10%	acceptable for general messages of low complexity
5-10%	good
<5%	excellent

The effect of background noise can also be taken into account using a second, more complex, formula. While very good correlation can be obtained between the direct to reverberant (D/R) ratio at 2000 Hz , the technique only tests the system over one frequency band. The method effectively assumes that the system is well behaved and controlled at other frequencies. It is therefore essential to carry out supplementary measurements such as overall frequency response and impulse response of D/R ratios at other frequencies, for example 500 Hz, 1000 Hz and 4000 Hz.

VI. CONCLUSIONS

A thorough acoustical study and speech intelligibility evaluation for mosques has to be carried out. Mosque is a challenging space for acoustics due to its volume, form and shape, and its main function for speech. Available literature on mosque's acoustics, thus far, reveals that the studies were initiated only when the mosque was problematic in terms of acoustics and speech intelligibility. In most cases, the problem is rectified only after the room is occupied. It would be better if the problem could be diagnosed at the initial design process since the rectification cost is expensive.

REFERENCES

[1] D. Davis and C. Davis. *Sound System Engineering*. 2nd Ed. Massachusetts: Focal Press, 1997.

[2] P. Mapp, "Reaching the audience," Sound and Video Contractors, Kansas: Intertec Publishing, Vol. 17, No. 11, pp. 17-32, October 1999.

[3] M. Barron. *Auditorium Acoustics and Architectural Design.* 2^{nd} Ed, London: E&FN Spon, 1998.

[4] Geoffrey K. Goacher, "A primer on listening room audio and acoustics," User's Guide for Perfecting Listening Room Audio and Acoustics with ETF 4.0, Ontario, Canada: , Acoustics Research Associates, March 1998.

[5] Donald E. Hall. *Basic Acoustics*, New York: Harper and Row, 1987.

[6] V. Peutz, "Speech reception and information," Syn Aud Tech Topics, Vol. 15, No. 12, 1978.

[7] D. Templeton, P. Sacre, P. Mapp, D. Saunders. Acoustics in the Built Environment, Oxford: Butterworth-Architecture, 1993.