December 2014

Speech Perception in Reverberated Condition By Cochlear Implants

Moulesh Bhandary
University of Wisconsin-Milwaukee

SPEECH PERCEPTION IN REVERBERATED CONDITION BY COCHLEAR IMPLANTS

by

Moulesh Bhandary

A Thesis Submitted in

Partial Fulfillment of the

Requirements for the Degree of

Master of Science

in Engineering

at

The University of Wisconsin – Milwaukee

December 2014

ABSTRACT

Speech Perception in Reverberated Condition by Cochlear Implants

by

Moulesh Bhandary

The University of Wisconsin- Milwaukee, 2014
Under the Supervision of Professor Yi Hu

Previous Studies for bilateral cochlear implants users examined cocktail –party setting under anechoic listening conditions. However in real world listeners always encounter problems of reverberation, which could significantly deteriorate speech intelligibility for all listeners, independent of their hearing status.

The object of this study is to investigate the effects of reverberation on the binaural benefits for speech recognition by bilateral cochlear-implant (CI) listeners.

Bilateral CI subject was tested under different reverberation conditions. IEEE recorded sentences from one male speaker mixed with either speech shaped noise (ssn), energy masking, or with 2 female competing takers (2fsn), informational masking, at different signal –noise –ratios (SSN) were used as stimuli. The male target speech was always set at 90˚ azimuth (from the front), while the masker were placed 0˚, 90˚, 180˚ azimuth (0˚ implied left, 180˚ implied right). Generated stimuli were presented to

Bilateral Cochlear Implant subjects via auxiliary input, which was connected to sound processor in a double wall sound attenuated booth. In each condition, subject was tested with individual ear alone, as well as with both ears.

Prior studies predict there would be decrease in speech intelligibility in reverberated condition as compared with anechoic environment. As predicted we saw a decrease in speech intelligibility in reverberated condition as compared with anechoic environment as reverberant environment produce more masking than the less reverberant environment do. We also observed that benefit of spatial hearing in reverberant environment. We observed that when the masking was placed at the better ear the subject performed better than the masking placed the other ear. We also observed the reverberation effect on energetic and informational masking. We observed that when the target and interfere are spatially separated, reverberation had greater detrimental effect on informational masking than energetic masking, and when the target and interfere were co-located the energetic masking results performed better than informational masking.

Due to time limitation and subject availability, test was done with one CI subject. Further testing and research on this topic, would help to understand the effect/s the informational masking vs energetic masking in reverberated conditions.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank Dr. Yi Hu, my advisor, for guiding and stimulating my interest in Cochlear Implant research. I would also like to thank him for giving a part of his busy schedule for simulating my interested in cochlear implants, with helping us in our research, acting as mentor and good colleague. I would take this opportunity to thanks the professors of CEAS for educating us and helping us during our difficult times and unknown solutions.

I would like to thank my fellow lab mates for simulating a relaxed surrounding in our lab.

Finally, I wish to thank my family, without their support I would never have got as far as I have.

Chapter 1: Introduction

1.1 Background:

A cochlear implant is a small, complex electronic device that can help to provide a sense of sound to a person who is profoundly deaf or severely hard-of-hearing. People with mild or moderate sensorineural hearing loss are generally not candidates for cochlear implantation. Their needs can often be met with hearing aids alone or hearing aids with an FM system. After the implant is put into place, sound no longer travels via the ear canal and middle ear but will be picked up by a microphone and sent through the device's speech processor to the implant's electrodes inside the cochlea. Thus, most candidates have been diagnosed with a severe or profound sensorineural hearing loss. Cochlear implants are designed to help severely to profoundly deaf adults and children who get little or no benefit from hearing aids. Even individuals with severe or profound "nerve deafness" may be able to benefit from cochlear implants. The presence of auditory nerve fibers is essential to the functioning of the device: if these are damaged to such an extent that they cannot receive electrical stimuli, the implant will not work.

A cochlear implant is very different from a hearing aid. Hearing aids amplify sounds so they may be detected by damaged ears. Cochlear implants bypass damaged portions of the ear and directly stimulate the auditory nerve. Signals generated by the implant are sent by way of the auditory nerve to the brain, which recognizes the signals as sound. Hearing through a cochlear implant is different from normal hearing and takes time to learn or relearn. However, it allows many people to recognize warning signals, understand other sounds in the environment, and enjoy a conversation in person or by telephone.

Post-lingually deaf adults, pre-lingually deaf children and post-lingually hard of hearing people (usually children) who have lost hearing due to diseases such as CMV and meningitis, form three distinct groups of potential users of cochlear implants with different needs and outcomes. Those who have lost their hearing as adults were the first group to find cochlear implants useful, in regaining some comprehension of speech and other sounds. The outcomes of individuals that have been deaf for a long period of time before implantation are sometimes astonishing, although more variable. Another group of customers are parents of children born deaf who want to ensure that their children grow up with good spoken language skills. The brain develops after birth and adapts its function to the sensory input; absence of this has functional consequences for the brain, and consequently congenitally deaf children who receive cochlear implants at a young age (less than 2 years) have better success with them than congenitally deaf children who first receive the implants at a later age, though the critical period for utilizing auditory information does not close completely until adolescence. The third group who will benefit substantially from cochlear implantation are post-lingual subjects who have lost hearing: a common cause is childhood meningitis. Young children (under five years) in these cases often make excellent progress after implantation because they have learned how to form sounds, and only need to learn how to interpret the new information in their brains

According to the Food and Drug Administration (FDA), as of December 2012, approximately 324,200 people worldwide have received implants. In the United States, roughly 58,000 adults and 38,000 children have received them. A cochlear implant costs approximately $60,000 (including the surgery, adjustments, and training).

In India, there are an estimated 1 million profoundly deaf children, only about 5,000 have cochlear implants (from Wikipedia).

A cochlear implant will not cure deafness, but is a prosthetic substitute for hearing. Some recipients find them very effective, others somewhat effective and some feel worse overall with the implant than without. For people already functional in spoken language who lose their hearing, cochlear implants can be a great help in restoring functional comprehension of speech, especially if they have only lost their hearing for a short time.

Individuals who have acquired deaf blindness (loss of hearing and vision combined) may find cochlear implants a radical improvement in their daily lives. It may provide them with more information for safety, communication, balance, orientation and mobility and promote interaction within their environment and with other people, reducing isolation. Having more auditory information than they may be familiar with may provide them with sensory information that will help them become more independent.

Many CI users describe initial sound after surgery as robotic sound of human voices, some decibel it as similar to radio static or voices as being cartoonish, though after a year with the implant users find it sound normal. Even modern cochlear implants have at most 22 electrodes to replace the 16,000 delicate hair cells that are used for normal hearing. However, the sound quality delivered by a cochlear implant is often good enough that many users do not have to rely on lip reading in quiet conditions. In noisy conditions however, speech understanding often remains poor

Many things determine the success of implantation. Some of them are:

- How long the patient has been deaf--as a group, patients who have been deaf for a short time do better than those who have been deaf a long time
- How old they were when they became deaf--whether they were deaf before they could speak

- How old they were when they got the cochlear implant--younger patients, as a group, do better than older patients who have been deaf for a long time
- How long they have used the implant
- How quickly they learn
- How good and dedicated their learning support structure is
- The health and structure of their cochlea--number of nerve (spiral ganglion) cells that they have
- Implanting variables, such as the depth and type of implanted electrode and signal processing technique
- Intelligence and communicativeness of patient

1.2 Parts of the cochlear implant

The implant is surgically placed under the skin behind the ear. The basic parts of the device include:

External:

- one or more microphones which picks up sound from the environment
- a speech processor which selectively filters sound to prioritize audible speech, splits the sound into channels and sends the electrical sound signals through a thin cable to the transmitter,
- a transmitter, which is a coil held in position by a magnet placed behind the external ear, and transmits power and the processed sound signals across the skin to the internal device by electromagnetic induction,

Internal:

A receiver and stimulator secured in bone beneath the skin, which converts the signals into electric impulses and sends them through an internal cable to electrodes,

- An array of up to 22 electrodes wound through the cochlea, which send the impulses to the nerves in the scala tympani and then directly to the brain through the auditory nerve system. There are 4 manufacturers for cochlear implants, and each one produces a different implant with a different number of electrodes. The number of channels is not a primary factor upon which a manufacturer is chosen; the signal processing algorithm is also another important block.

A cochlear implant receives sound from the outside environment, processes it, and sends small electric currents near the auditory nerve. These electric currents activate the nerve, which then sends a signal to the brain. The brain learns to recognize this signal and the person experiences this as "hearing".

The cochlear implant somewhat simulates natural hearing, where sound creates an electric current that stimulates the auditory nerve. However, the result is not the same as normal hearing.

The implant consists of an external portion that sits behind the ear and a second portion that is surgically placed under the skin (see figure 1). An implant has the following parts:

- A microphone, which picks up sound from the environment.
- A speech processor, which selects and arranges sounds picked up by the microphone.
- A transmitter and receiver/stimulator, which receive signals from the speech processor and convert them into electric impulses.
- An electrode array, which is a group of electrodes that collects the impulses from the stimulator and sends them to different regions of the auditory nerve.

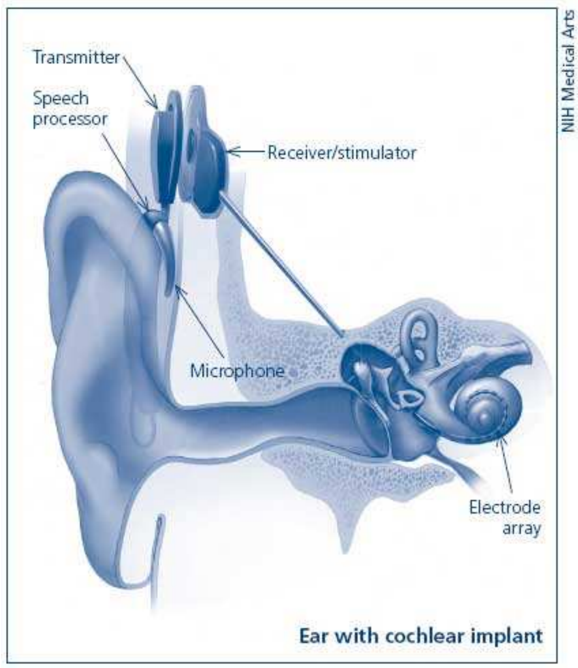


Figure 1: Ear with Cochlear Implant, Credit: NIH Medical Art

Currently (as of 2013), the three cochlear implant devices approved for use in the U.S. are manufactured by Cochlear Limited (Australia), Advanced Bionics (USA, a division of Sonova) and MED-EL (Austria). In Europe, Africa, Asia, South America, and Canada, an additional device manufactured by Neurelec (France) is available. Lastly, a device made by Nurotron (China) is available in some parts of the world. Each manufacturer has adapted some of the successful innovations of the other companies to its own devices

1.3 Main Problems Faced By CI users

1. speech recognition with cochlear implants
2. implant user can talk on the phone in a quiet environment
3. Listening in Echo
4. Listening in Reverb
5. Speech perception and localization with adults with bilateral sequential cochlear implants
6. Music perception with cochlear implants

2 Reverberation

A Reverb simulates the component of sound that results from reflections from surrounding walls or objects. It is in effect a room simulator. Some people think it's just a delay effect with some filters, but its way more complex than that. Reverb effects (software plug-in or external hardware units) provide an interface to their changeable parameters that need some explaining. Let's look at a simple room first.

Reverberation is the collection of reflected sounds from the surfaces in an enclosure like an auditorium. if it is excessive, it makes the sounds run together with loss of articulation - the sound becomes muddy, garbled. To quantitatively characterize the reverberation, the parameter called the reverberation time is used

2.1 Basic Simulation of a Room

Our model is a simple room with four straight walls, a sound source and a listener. In Figure 2 the arrows stand for the path of traveling sound.

The listener hears the DIRECT signal first. The DIRECT signal is also referred to as the DRY part of the signal when using any effect. Most digital reverbs produce two parts: The Early Reflections and the Reverb component.

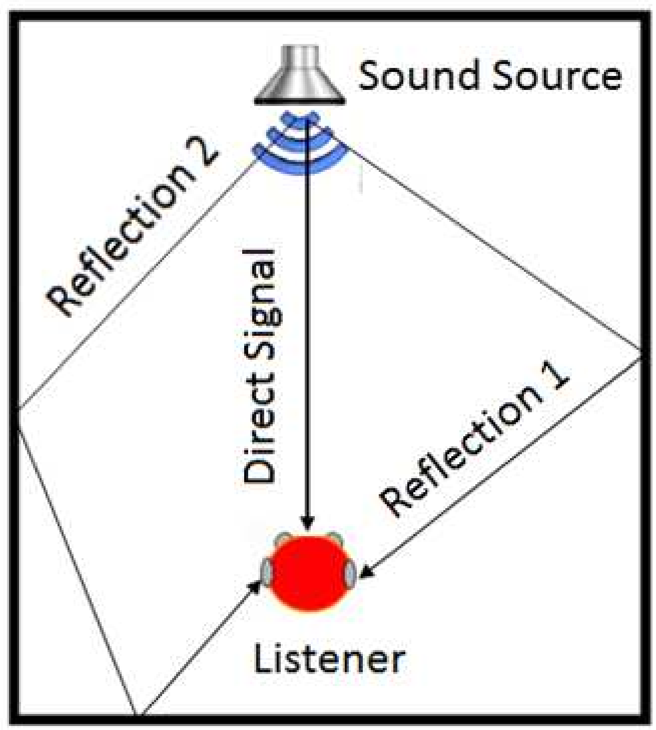


Figure 2: Reverberation

Early Reflections

The first Early Reflection reaches the listener milliseconds after the direct signal does. The path of the Early Reflections is longer. The difference in time between the arrival of the direct signal and the first Early Reflections is measured in milliseconds.

The sound reflects off the walls and objects in the room, and in time individual reflections disappear and the Reverb develops.

Predelay

The time between the reception of the DIRECT signal by the listener and start of the Reverb portion of the effect is called Predelay. This is a parameter in many digital reverb effects, and it is expressed in milliseconds (ms).

2.2 Reverb Time

The time difference between switching off any sound generator and the level of the reverb resulting from that sound dropping by 60dB is called RT60.

This is usually referred to as the Reverb Time. When anyone refers to the reverb time of a real room or that of a digital reverb, RT60 is what they're talking about.

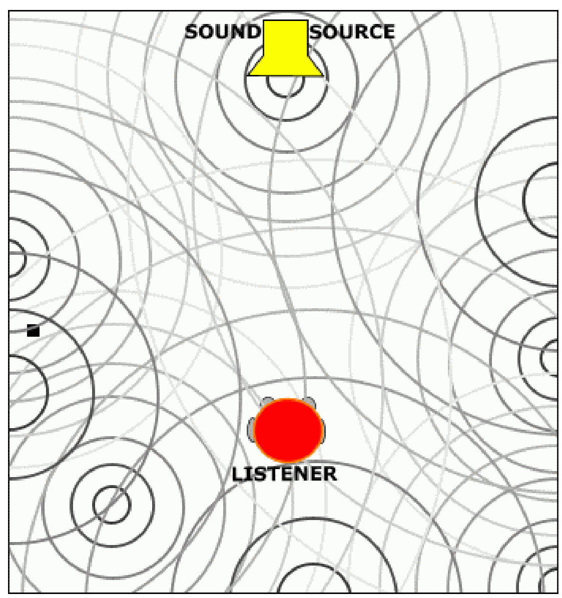


Figure 3: Reverberation Condition in Room

Most digital reverbs feature this as a parameter.

2.3 Problems with Reverberation

Reverberation can cause significant deterioration in speech intelligibility. Human ears or microphones are susceptible to reverberation from voice sources. Reverberation is a common phenomenon in enclosed spaces.

Several Researchers have noted that detrimental effects of reverberation time (RT60) on speech perception[i]. It is suggested that reverberation flattens format transaction in vowels, resulting in weak- energy speech units being masked by preceding segments with strong energies. This causes smears in spectral cues, reduction in temporal amplitude modification and thus increases low frequency energies which thereby cause masking of higher frequency components[ii]

Its well established that normal-hearing (NH) listeners have a remarkable ability to perceptually segregate competing voices from the target voice amid a background, a formidable task that has been termed the “cocktail-party” problem e.g., Cherry, 1953. When the target voice and the interfering voices are spatially separated NH listeners are able to take advantage of the favorable SNR (Signal to Noise Ratio) at the “better “ear due to head shadow effect.

NH listeners are able to exploit a number of cues that help them cope with the cocktail party problem. In addition, NH Listener’s are able to receive binaural advantage resulting from binaural unmasking in low frequencies.(Bronkosrt and Pomp, 1988; Zurek 1993)

A lot of researches have been done to understand perceptual process used by NH listener’s to segregate a target voice from competing background noises. The objective of this study is to investigate the effects of reverberation on the binaural benefits for speech recognition by bilateral cochlear-implant (CI) listeners. Much research needs to be done to help CI users perform better in reverberated surrounding like churches, conference rooms where noise may be present from surrounding.

A number of studies have been done on with CI users where the target and masker were coincident or spatially separated (e.g. Litovsky et al, 2006). In the study Tyler et al (2002) data from nine CI subjects, who had bilateral implant 3 months prior to the test, results showed that when the noise was spatially separated from target voice, the subjects showed a significant head shadow advantage but few subjects showed binaural-interaction benefit arising from using both ears over better ear with better SNR. Similar test results were published by Muller et al (2002) where speech was presented from front and steady speech –shaped noise was present at either +90 degree or -90 degree azimuth at fixed SNR (10dB). Their results indicated significant head-shadow benefit as well as small binaural- interaction benefit.

Chapter 3 Head Related Transfer Function (HRTF)

Binaural hearing is ability helps human and animal to judge the direction of the sound source. Using the two ears, humans have been able to localize the sound sources.

Lord Rayleigh (John William Strutt) (during 1877-1878), is named to be the founder of localization process. He noted that if a sound source is in the ipsi-lateral ear (on the same side) , then the head makes a shadow cast in the contra-lateral ear. This makes the signal in the contra-lateral ear more attenuated than ipsi-lateral ear. He also noted that different parameters affect the localization at low and high frequencies. His theory is named as "Duplex Theory". Many models of Binaural processing were created over the last century, some of them are listed below

- "Spherical Head Model" – Lord Rayleigh, 1907 and Woodworth/Scholsberg,1954,
- "Direct Cross-correlation of stimuli model" – Sayers and Cherry 1957
- "The Binaural cross-correlation model" – Jeffress 1956,
- "Direct Comparison of amount of left sided and right sided internal response stimuli model" - Bergerijk 1962
- "Interaural comparison auditory –nerve activity model" –Colburn 1973-1977

3.1. Binaural Perception

3.1.1 Binaural Cues

There are two important binaural physical cues in the horizontal plane.

a) Interaural Time Difference (ITD), delays

b) Interaural Level Difference (ILD), intensity

3.1.1. A) Interaural Time Difference (ITD), delays

The sound source arrives at different times in ipsi-lateral and contra-lateral ear is called ITD. ITD is dominant cue at frequencies lower than 1500 Hz. The wavelengths of frequencies lower than 1500 Hz are comparable with human size head. The minimum ITD is zero and maximum ITD is about 600-800 µs. ITD is more sensitive in near field (less that 1 meter source distance) than in far-field.

Using a simple single sound source at azimuth ‘θ’ and spherical head model of radius ‘a’, ITD can be obtained using Rayleigh Spherical Head Model, with sound source at Infinity.

$$\mathrm{ITD} = \left|\frac{a}{c}\,(\theta + \sin(\theta))\right| \qquad -\pi/2 \le \theta \le \pi/2 \qquad \ldots\ldots\ldots\; 3.1$$

where c is the speed of sound and θ is the azimuth angle between center of head and azimuth plane.

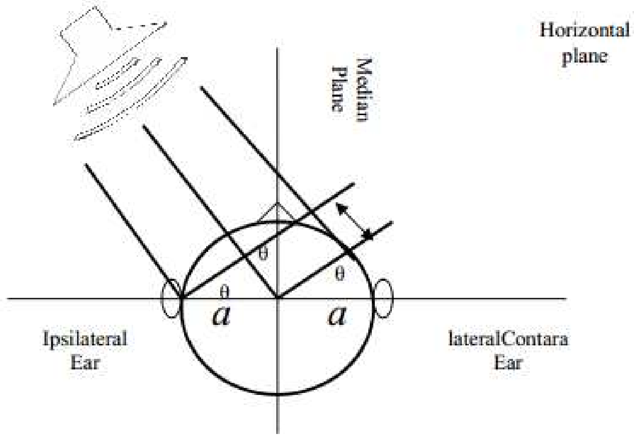


Figure 4: Spherical Model in Horizontal Plane.

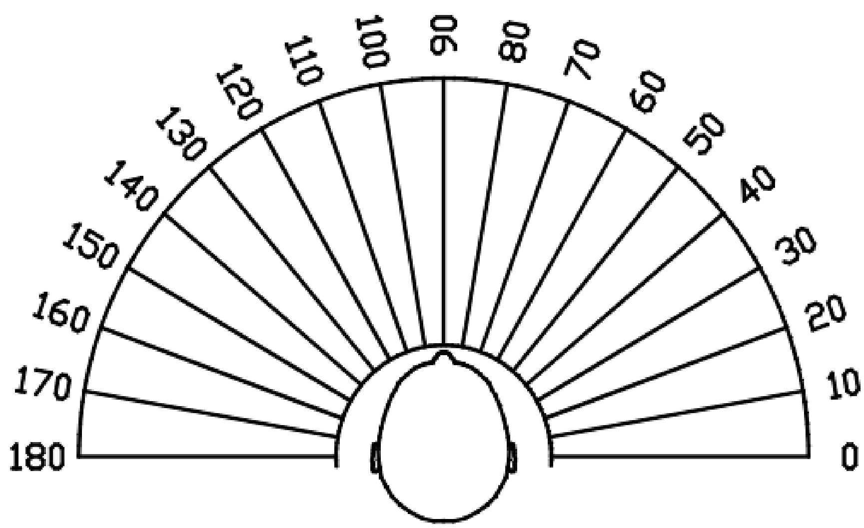


Figure 5 : Semi Circle of Horizontal Plane with 90 ° in front of person.

Using the equation 3.1, we can calculate that

ITD = 0, when sound source is in front of head and

ITD = 1.57 (a/c), when sound source is located at one of the two ear

The above equation is frequency independent, but in some models ITD is dependent on frequency.

3.1.1. B) Interaural Level Difference (ILD), delays

The sound Pressure level difference between the ipsi-lateral and contra-lateral ear is called ILD. ILD is a dominant cue at frequencies higher than 1500 Hz. ILD occurs because of the head shadow cast in the collateral ear. ILD dependencies to frequency are shown in the figure below.

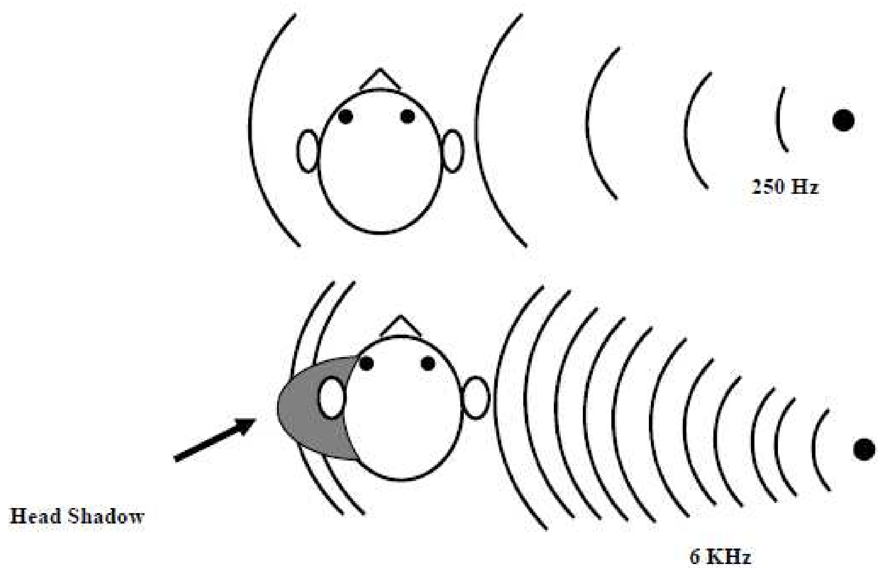


Figure 6 : Semi Circle of Horizontal Plane with 90 ° in front of person.

ILD is nonlinear with frequency and is strongly dependent on frequency over audible spectrum sound waves because more sound waves are scattered as the head diameter increases. The wavelength and diffraction also increase rapidly as frequency increases.

As seen noted by most research papers, smallest detectable ILD is 0.5 dB, regardless of frequency. The far-field ILD doesn't exceed 5-6 dB where as near field ILD exceeds 15 dB at 500 Hz[iii].

3.1.2 Head Related Transfer Function

A Head Related Transfer function (HRTF) is a response that characterizes how an ear receives a sound from a point in space. A pair of HRTFs for two ears' can be used to synthesize a binaural sound that seems to come from a particular point in space. HRTF is transfer function describing how a sound from a particular in space will arrive at the ear, generally outer ear of auditory canal. It depends on Frequency and azimuth in 2D space. Far field HRRTF is attenuated inversely by range where as near field follows ILDS changes.

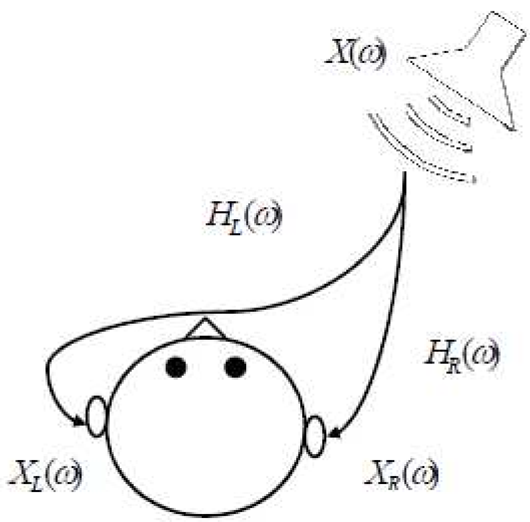


Figure 7 : HRTF Filtering effect on left & Right Ear

Signals received by the two ears are as follows

Left Ear $X_L(w) = H_L(w) . X(w)$

Right Ear $X_R(w) = H_R(w) . X(w)$

$H_L(w)$ and $H_R(w)$ are the frequency response of transformation for left and right ears respectively. $X_L(w)$ and $X_R(w)$ are signals received on left ear and right ear respectively. $X(w)$ is signal as shown in figure above. Dot (.) implies convolution.

In this research Aachen Impulse Response (AIR) database [v] is used to generate the required stimuli for left ear and right ear. Air Database is a set of impulse response that were measured in variety of rooms, meeting room, lecture room, stairway, corridor, aula carolina. The version of Air Database used for this research, uses the binaural room impulse response (BRIR) measured with a dummy head in different location with different acoustical prosperities, such as reverberation time and room volume. All the impulse responses of Air Database are stored as double precision binary floating-point MAT-files. Convolving the required .mat files with the sound source and noise conditions at specified SNR, the required stimuli was obtained.

3.1.3 Minimum Audible Angle

The just noticeable difference in Azimuth perceptible by listener is measure using the Minimum Audible Angle plot as show below. Although dependent on both individual, type of sound, nature of environment “Ambience”; under ideal conditions most listeners’ can detect change in angle of one degree when the source is straight ahead. This accuracy drops off as the source moves to the side of the head or in the case of pure tones, when the frequency lies between 1500 and 2000 Hz.

Mills, in 1958[iv], is credited to obtaining the MAA (Minimum Audible Angle) as function of Frequency and Azimuth. 1 Degree MAA is proportional to smallest detectable ITD, about 10 µs. As frequency increases MAA also increases. MAA is symmetric around 90° in spherical head model.

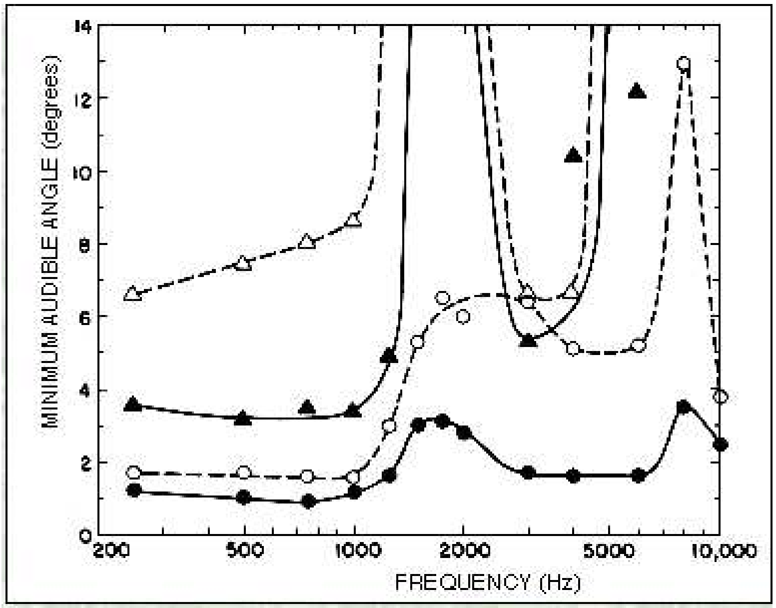


Figure 8: Minimum Audible Angle

Chapter 4: Testing and Conclusions

Previous studies have examined speech recognition by bilateral cochlear implant users in cocktail –party setting under anechoic listening condition. However in real world listening conditions, the speech stimuli is mixed with not just noise, subjects always encounter problems of reverberation. Reverberated speech deteriorates speech intelligibility for all listeners. In this study we studied the effect of reverberation by bilateral cochlear implant user. The interaction between masker types, spatial location and degree of reverberation will be discussed.

4.1 Subject and Speech Stimuli.

Post – Lingual deafened adults, wearing bilateral Cochlear Implant (CI) users were recruited for this testing. The Testing of subjects was conducted at UW – Milwaukee. All the subjects recruited were native speakers of American-English language and were paid an hourly wage for their participation. All subjects had a minimum of one year experience using their implant device and they used their own device while testing. The speech stimuli used for testing were from IEEE (Institute of Electrical and Electronic Engineers) database (IEEE, 1969). A male talker was recruited to record the IEEE database, which has 72 lists of 10 sentences each. The rootymean square (RMS) value of all the sentences was equalized to the same value corresponding to 64 dB

Aachen Impulse Response (AIR) Database[v] is used to generate the HRTF of selected room. Air Database is set of impulse responses recorded in wide variety of rooms, which allows its users to simulate realistic models in reverberated environments with a special focus on hearing aid applications. For our testing purpose we selected the binaural room impulse response in the staircase and in the Aula Carolina Aachen, with the dummy head. Aula Carolina Aachen is the former church in Aachen, Germany with ground area of 570 m^2 and with a high ceiling showing a very strong reverberation effect.

Figure 10: Staircase Recording

Figure 9 : Aula Carolina Recording

Since AIR database included the BRIR's with various azimuth angles between the head and desired source for staircase and aula carolina, for various distance from source, these two rooms were selected to generate HRTF. BRIR (Binaural Room Impulse Response) were generated using the dummy head option at different locations; so different stimuli would be generated with different acoustical properties such as reverberation time and room volume. This database allowed us to investigate the head related room response transfer functions for the 2 rooms for different azimuth angles for various distance from source..

In order to generate the stimuli for the study, the HRTF's were obtained from each reverberation condition convolved with the signal files in MATLAB. Signal files were either files from IEEE test materials or noise maskers. Two different noise maskers were used for this study to study the effect of energetic masking vs informational masking in reverberated condition. Speech shaped noise was used as energetic masking, while two female competing speeches was used as informational masking. The male target speech was always set at 90˚ azimuth (from the front), while the masker were placed 0˚, 90˚, 180˚ azimuth (0˚ implied left, 180˚ implied right). Generated stimuli were presented to Bilateral CI subject via auxiliary input.

In the studies of reverberated speech on CI[vi], it was shown that late reverberation was more detrimental to speech than early reflection to CI subjects. In the present study we use discrete- time domain to investigate the reverberation perception by CI subjects. Let s[n] denote the clean discrete-time speech signal , h[n] denote the HRIR for specified distance from source and set azimuth, n[n] denote the noise signal, then the reverberated stimuli is obtained by $x[n] = s[n] * h[n] + n[n] * h[n]$, (4.1)

where * indicates the discrete-time convolution operator.

The casual HRIR filter h[n] can be decomposed into three components

- h[0] represents the direct path,
- $h_e[n]$ represents the early reflection
- $h_l[n]$ represents the late refection

A simplified version of statistical model for the room impulse response filter in Polack (1988) can be described as random process with an exponentially decayed envelope signal

$$h[n] = \begin{cases} 0, for\ n < 0 \\ h[0], for\ n = 0 \\ h_e[n], for\ 1 \le n \le T_e.f_s - 1 \end{cases}$$

$$h[n] = \begin{cases} 0, for\ n < 0 \\ h[0], for\ n = 0 \\ \mu[n]\ e^{-v\frac{n}{f_s}}, for\ 1 \le n \le T_e.f_s - 1 \end{cases}$$

where f_s denotes the sampling frequency, Te denotes duration for early refection, v denotes RT60 , reverberation time T60. $\mu[n]$ represents random variable sequence of independent and identical normal distribution.

The Reverberation time is denoted as $v = \frac{3\ln(10)}{T_{60}}$.

Since the noise (masker) and speech were placed at different azimuth, equation (4.1) can be further decomposed as

$$s_L[n] = s[n] * h_{1L}[n] + n[n] * h_{2L}[n] \qquad (4.2)$$

$$s_R[n] = s[n] * h_{1R}[n] + n[n] * h_{2R}[n] \qquad (4.3)$$

stimuli= [s_L[n] s_R[n]] (4.4)

where $s_L[n]$ and $s_R[n]$ represents the stimuli on the left ear and right ear respectively

$s[n]$ and $n[n]$ represent the speech and noise(masker) to produce the required stimuli.

$h_{1L}[n]$, $h_{1R}[n]$, $h_{2L}[n]$ and $h_{2R}[n]$ represents the room impulse response for speech on left channel for azimuth 1, for speech on right channel for azimuth 1, for noise (masker) on left channel for azimuth 2 and for noise(masker) on right channel for azimuth 2 respectively. As stated before azimuth 1 for speech is always 90º (from front), while azimuth 2 for noise (masker) can be 0º, 90º, or 180º.

With the implementation of the algorithm described in section B of the Appendix, stimuli were generated for two different kinds of noise (masker) at different azimuths, at different signal to noise ratios.

4.2 Procedure and Testing

All stimuli's were presented to CI listener directly through the CI device audio cable, which was connected to a processing unit. Auxiliary input jack of the CI device was connected to sound processor in a double wall sound attenuated booth. Prior to testing each subject participated in a short practice session to gain familiarity with the listening task. Participant's signature (consent) was obtained on institutional review board approval forms and consent forms before testing commenced. During testing to avoid fatigue, subjects were given breaks after 2-3 conditions.

In this testing each Subject participated in

a) Reverberated (speech +speech shaped noise) for SNR at 5dB for Stairs at 1m and Aula Carolina at 3m

b) Reverberated (speech +speech shaped -noise) for SNR at 10dB for Stairs at 1m and Aula Carolina at 3m

c) Reverberated (speech +2 Female talker- noise) for SNR at 5dB for Stairs at 1m and Aula Carolina at 3m

d) Reverberated (speech +speech shaped noise) for SNR at 10dB for Stairs at 1m and Aula Carolina at 3m

e) (Speech + noise) for SNR at 5 dB, 10 dB

f) Clean speech

Two IEEE list were used per condition none of the list were repeated across different conditions. During the testing, the participants were given blank answer sheet to write on, corresponding to stimuli list and were allowed to repeat the sentence only once. The Participants would try to identify as many words as they could identify when the stimuli was played and wrote them corresponding to the sentence number of that particular list stimuli. The responses of each participant were scored off line based on the number of words correctly identified. All words of IEEE list were scored. Finally Percent correct score for each condition was calculated by dividing the correct number of words by total number of words in the particular list.

4.3 Results

Table 1 : Results for only Speech +Noise Stimuli for Subject 1

Noise Type	2fsn (2 Female competing Talker)		
		Left ear Only	Right ear only
5dB		55.68%	6.17%
10dB		74.07%	32.98%
Noise Type	**SSN (Speech Shaped Noise)**		
		Left ear Only	Right ear only
5dB		61.90%	7.41%
10dB		72.22%	28.92%

From Table 1, we see that subject 1 has right ear cochlear implant dominant over the left ear cochlear implant. Table 2 shows the result obtained with speech shaped noise masker.

Table 2 : Results Subject 1, at 5db SNR , Noise Type: SSN

	5 dB		
Noise Type		Noise -SSN	
Noise Angle	0	90	180
Aula -3m	54.55%	38.41%	8.07%
List Used	3-4	5-6	7-8
Stair - 1 m	89.31%	48.13%	29.56%
List Used	27-28	29-30	31-32

Table 3 : Results Subject 1, at 10 db SNR , Noise Type: SSN

	10 dB		
Noise Type	Noise -SSN		
Noise Angle	0	90	180
Aula -3m	41.89%	38.93%	17.61%
List Used	9-10	11-12	13-14
Stair - 1 m	90.13%	71.60%	43.40%
List Used	33-34	35-36	37-38

Table 4 : Results Subject 1, at 5 dB SNR , Noise Type: 2FSN

	5 dB		
	Noise -2fsn		
Noise Angle	0	90	180
Aula -3m	25.00%	22.01%	1.32%
List Used	15-16	17-18	19-20
Stair - 1 m	52.56%	60.90%	12.24%
List Used	39-40	41-42	43-44

Table 5 : Results Subject1, at 10 dB SNR , Noise Type: 2FSN

		10 dB	
Noise Type		Noise -2fsn	
Noise Angle	0	90	180
Aula -3m	48.15%	29.75%	9.74%
List Used	21-22	23-24	25-26
Stair - 1 m	47.98%	47.83%	38.65%
List Used	45-46	47-48	49-50

Table 6: Chart Room Condition: Stair case

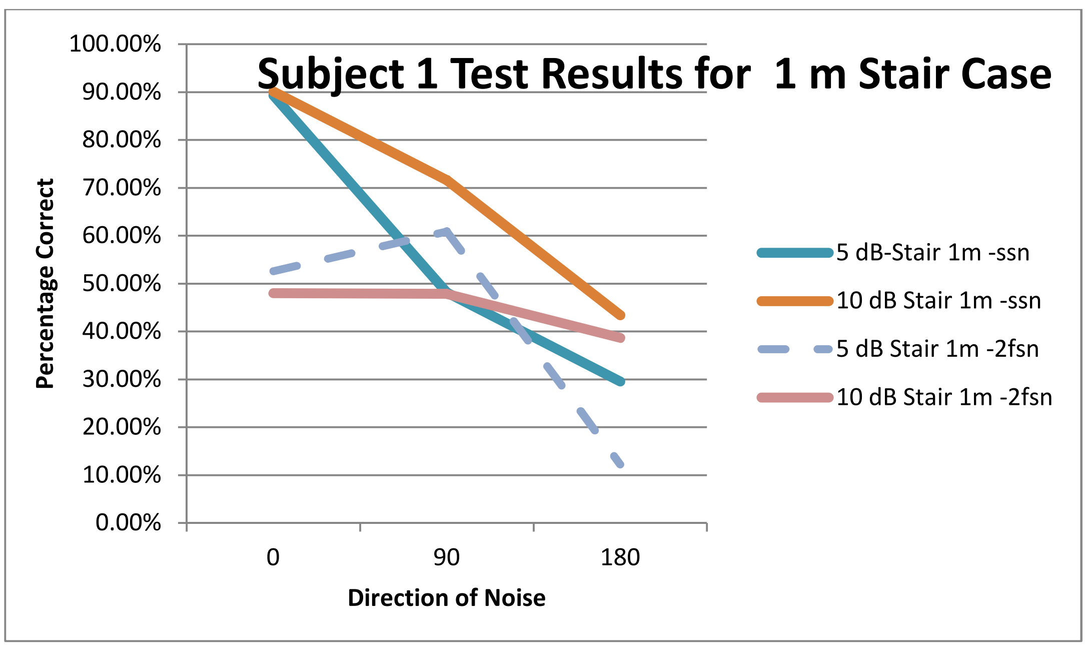
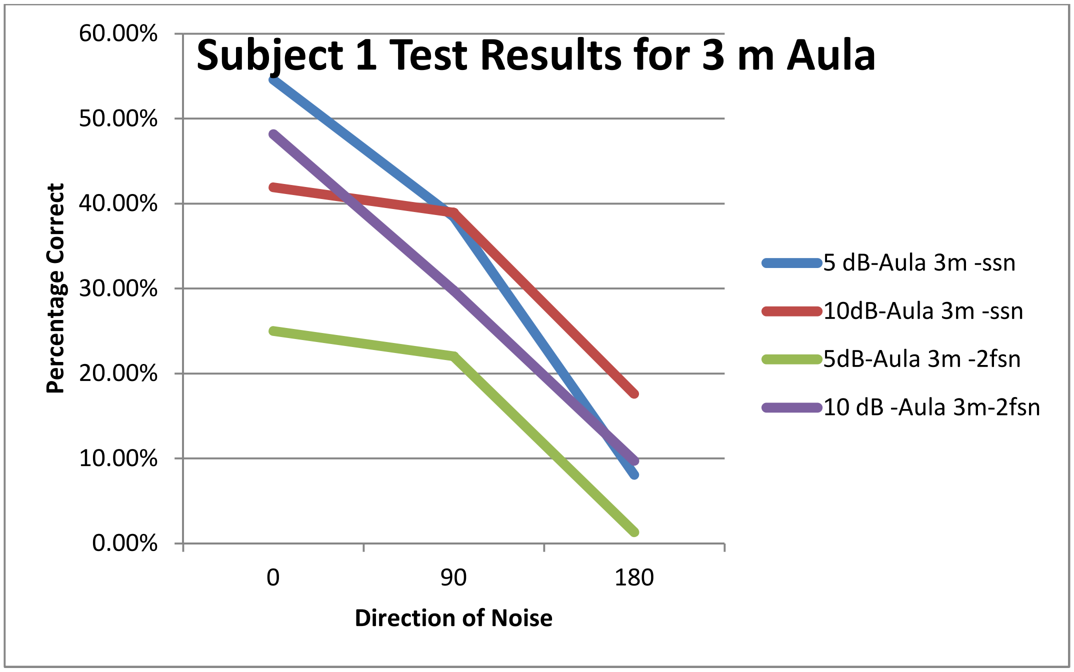


Table 7 : Chart Room Condition: Aula



4.4 Discussion and conclusion.

Results from table 2 indicate that subject 1, uses their left cochlear implant better than the right cochlear implant, so the dominant ear for this CI subject is their left ear. The stair case for 1m has lower reverberation than aula at 3m, we can conclude from table 2 to 5 that as reverberation increase the speech intelligibility of listener decrease, which is in par with many of the earlier studies done. We can also predict that there is a strong and negative relationship between speech perception and amount of acoustical reverberation.

For 5dB SNR condition, the intelligibility score were around 55% for SSN type noise to 25% for 2FSN type noise at 0 °, 38% for SSN type noise to 22 % for 2FSN type noise at 90 °, and 8% for SSN type noise to 1% for 2FSN for 180°, which could imply that at 5 dB SNR SSN dominates 2FSN, which implies that energy masking dominates informational masking at 5dB SNR. 10 dB can be considered as the ceiling effect of this subject. Also we can see the benefit of spatial hearing in reverberant condition. Since Left ear of the subject 1 is dominant, at 5dB the subject takes advantage of the Head Shadow effect which boosts the hearing when the noise is placed either left of right of target speech. Intelligibility of speech is reduced when speech and noise are placed in the same direction, 90° degree i.e from the front.

Further testing with bilateral CI can help boost the confidence in this result. In this testing we tried to find an interaction between masker types, spatial location and degree of reverberation. We can hypothesize from the result that as reverberant environment decreases the intelligibility of CI users than an anechoic room, since Reverberant Environment Produced more masking than less reverberant environment. At 5 dB we can suggest that energy masking dominates informational masking and also spatial separation between noise and speech boost the speech Intelligibility. This difference in performance helps us understand the performance benefit of the two ears that negatively affect benefit in bilateral CI under reverberant listening condition.

BIBLIOGRAPHY

[i] N´abˇelek and Picket 1974; N´abˇelek and Letowski 1985, 1988; N´abˇelek et al. 1989;
[ii] Bolt and MacDonald 1949; N´abˇelek and Letowski 1988; N´abˇelek et al. 1989; Assmann and Summerfield 2004
[iii] BRUNGART D. S. AND RABINOWITZ W. M., 1999 "AUDITORY LOCALIZATION OF NEARBY SOURCES. HEAD RELATED TRANSFER FUNCTIONS ",ACOUSTICAL SOCIETY OF AMERICA.
[iv] The moving minimum audible angle is smaller during self motion than during source motion W. Owen Brimijoin* and Michael A. Akeroyd
[v] Marco J, Magnus S, Peter V, "A Binaural Room Impulse Response Database for evaluation of Dereverberation Algorithms " Proc. Int. Conf. Digital Signal Process. (DSP), 2009
[vi] Yi H, Kosta K, "Effect of early and late reflections on intelligibility of reverberated speech by cochlear Implant Listners", J. Acoust . Soc. Am , Jan 2014

http://en.wikipedia.org/wiki/Cochlear_implant

http://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/ImplantsandProsthetics/CochlearImplants/ucm133345.htm

http://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/ImplantsandProsthetics/CochlearImplants/

http://report.nih.gov/NIHfactsheets/ViewFactSheet.aspx?csid=83

APPENDIX

A. List of Symbols

CI: Cochlear Implant

RIR: Room impulse Resposne

BRIR: Binaural Room Impulse Response

HRTF: Head Related Transfer Function

B. MATLAB CODE

1. To Calculate the Stimuli

```
%%This file generates stimuli for 1 noise ( 1 noise direction)
clear all;
close all;
clc;

currentfolder = pwd;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%% CHANGES ONLY TO MADE BelOw as
Mentione %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%(A)  Distance From Sources to be used in Air Database
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%% A) Room
%airpar.room = 11;   % aula_carolina 0:45:180
%airpar.room = 5;   % Staircase   0:15:180
%%airpar.room = 4;   % Lecture
```

```
room =11;                                     %% Put in Room Type
%%% only Stair case can be done here.
%%% For Staircase  Stairway: {1m, 2m, 3m}  %%% 1 ==> 1M, 2 ==> 2M ,3 ==>
3M
%% put the Distance Required
%%% For AULa     (1m, 2m, 3m, 5m,10m, 15m, 20m)

d=3;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%% TEST FOLDER NAME %%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

TestFolderName= 'SubjectName-Aula-Date';   %%%% Put the Name of
outputFolder
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%% Noise File Type %%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
noisefile = '2fsn-11062014-25k.wav';   %% Noise

%%% Path of Matlab Folder
pathh='C:\Users\bhandary\Desktop\Testing';
%% CHANGES ONLY TO MADE ABOVE NONE BELOW %%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%


%% First Run Trial_RoomIR.m then run this, Conv_HRTF_Audio_Script
disp('What list to start from ? ');
prompt = 'List No:';
result = input(prompt);
listNo = result;
if (listNo>=73)
    error('No Such List Number');
elseif (listNo==72)
    disp('list 72 and list 3 would be used in this process');
    disp('Type 0 to end process, type 1 to continue ');
    prompt = '0 or 1:';
    result = input(prompt);
    if (result ==0)
        error('User Prompted to cancel');
    end
end
prompt = 'What is the required Noise SNR ? ';
result = input(prompt);
nsnr = result;
currentlocation =pwd;
disp(' ')
disp('Direction of noise: (0° left, 90° frontal, 180° right) ');
prompt = 'What is the direction of Noise ? ';
result = input(prompt);
azimuth_noise = result;
a_n=num2str(azimuth_noise);
```

```
path_concat=strcat(pathh,'\','Database\IEEE\CleanVoice');
inpath_speech = path_concat;
path_concat = strcat(pathh,'\','Database\noise');
inpath_noise=path_concat;
path_concat = strcat(pathh,'\','recycle');
DDelete=path_concat ;

path_concat = strcat(pathh,'\','Reverb\TestMaterial\',TestFolderName);
output_folder= path_concat ; %OutputFolder
if (exist( output_folder, 'dir')~= 7)
    if ~(mkdir( output_folder))
        error( 'Cannot create output directory');
    end
end
path_concat = strcat(pathh,'\','OutPutFolder\Reverb\TestMaterial\',...
    TestFolderName,'-noise');   %% Noise foldername
output_foldernoise=path_concat;
if (exist( output_foldernoise, 'dir')~= 7)
    if ~(mkdir( output_foldernoise))
        error( 'Cannot create output directory');
    end
end


fspeech= fopen( strcat( output_folder, '\', 'allConds.txt'), 'at');
fnoise= fopen( strcat(output_foldernoise, '\', 'noiseConds.txt'), 'at');

fprintf(fnoise,'%s\n','----noise file ------' ,'noise
type :',noisefile);

noisefile= strcat( inpath_noise, '\', noisefile);

%%  this one does per list of 10 sentences per list
for l=listNo:1:listNo+1
    for i=1:1:10

        Speech=['S_' num2str(l) '_'  num2str(i) '.wav'];
        infile= strcat( inpath_speech, '\', Speech);
        noiseout=['n_' num2str(l) '_' num2str(i) '.wav'];
        outfile= strcat( output_foldernoise, '\', noiseout);
        nf=64;  %nf : normalize to nf dB
        m_addnoise(infile,noisefile, nsnr,nf,outfile);
        fprintf(fnoise,'%s\n\n',strcat('lists: ' ,num2str(l),'-
' ,num2str(l+1),...
            ' : ',noiseout,' Snr: ' ,num2str( nsnr),'db/','noise
direction:' ,a_n));
        [Y,targetSrate]=wavread(outfile);

    end
end
fclose(fnoise);


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%% Air Database
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%----------------------------------------------------------------------
----
% Load room impulse responses from the AIR database
%----------------------------------------------------------------------
----
% Details of the measured room impulse responses can be found in the
% corresponding papers:
%
% M. Jeub, M. Schaefer, and P. Vary
% "A Binaural Room Impulse Response Database for the Evaluation of
% Dereverberation Algorithms", in Proc. of 16th International
Conference on
% Digital Signal Processing (DSP), Santorini, Greece, 2009
%----------------------------------------------------------------------
---
path2output=output_folder;

[h_aula_L,h_aulanoise_L,h_aula_R, h_aulanoise_R, ...
    air_aula_L,air_aulanoise_L, ...
    air_aula_R,air_aulanoise_R,fig1,fig2]=HRTF_room(azimuth_noise,...
    targetSrate,d,room,path2output);

mk_folder=1;
for l=listNo:1:listNo+1
    if (l== 73)
        l=3
        disp('Reached end of list starting from list 3')
    else
        l=l
    end
    mk_folder= mk_folder+1;
    if mod(mk_folder,2)==0
        fprintf(fspeech,'%s\n\n','----Speech Stimuli ------',
strcat(date ,....
            '  \\noise type :',noisefile,...
            ' \\ Snr: ' ,num2str( nsnr),' db',' \\ noise
direction:' ,...
            num2str(air_aulanoise_R.angle),...
            ' degrees, \\ Room: ',air_aula_R.room,...
            ', \\distance from speaker:
',num2str(air_aula_R.distance), ....
            ' m','\\List Used :',num2str(listNo),'-
',num2str(listNo+1)));  %%for aula
    else
    end


    if mod(mk_folder,2)==0
        Stimuli_fol= strcat(output_folder, '\',num2str( l),'-
',num2str( l+1));
        mkdir(Stimuli_fol);
```

```
        fstimuli= fopen( strcat( Stimuli_fol, '\', 'StimuliConds.txt'),
'wt');
        fwavefileProp=fopen( strcat(Stimuli_fol, '\',
'wavefileProperties.txt'), 'wt');
        fprintf(fstimuli,'%s\n\n','----Speech Stimuli ------' ,....
        strcat('\\noise type :',noisefile,'\\ Snr:
' ,num2str( nsnr),'db/',....
        '\\noise direction:' ,num2str(air_aulanoise_L.angle),
'degrees,\\ Room: ',air_aula_R.room,...
        ',\\ distance from speaker: ',num2str(air_aula_R.distance), '
m'));  %%for aula

fprintf(fstimuli,'%s\n\n','version',num2str(air_aula_L.version),....
        'Head (Yes ==>1, No ==> 0) :=
',num2str(air_aula_L.head),'Distance :',...
         num2str(air_aula_L.distance),'Angle ',
num2str(air_aulanoise_L.angle),....
         'Misc :', num2str(air_aula_L.misc),'Microphone :' ,
num2str(air_aula_L.microphone),....
         'Left Channel is 1? ', num2str(air_aula_L.channel),'Right
Channel is 0? ',...
         num2str(air_aula_R.channel),....
         'Excitation', num2str(air_aula_L.excitation),' ' );
        movefile(fig1,Stimuli_fol);
        movefile(fig2,Stimuli_fol);
    else
    end
    for i=1:1:10
        speechfile=['S_' num2str(l) '_'  num2str(i) '.wav'];
        speechininput= strcat( inpath_speech, '\', speechfile);
        noiseinfile=['n_' num2str(l) '_' num2str(i) '.wav'];
        noiseinput = strcat( output_foldernoise, '\', noiseinfile);
        Stimu=['St_'  num2str(l) '_'  num2str(i) '.wav'];
        Stimuliout=strcat(Stimuli_fol, '\', Stimu);

        wReverbLeft =['ReverbLeft_' num2str(i) '.wav'];
        wReverbRight =['ReverbRight_' num2str(i) '.wav'];
        [Y,fs,nbit]=wavread(speechininput);
        [Yn,Fsn,nnbit]=wavread(noiseinput);
        if  (fs/Fsn)==1
            Yn =Yn;
        else
            [P,Q]=rat(fs/Fsn);
            Ynew =resample(Yn,P,Q);
            Yn=Ynew;
            error('Sampling rates of noise and Speech donot match');
        end

        %% Get the same Yn as Y,

        if (length(Y)==length(Yn))
            Yn= Yn;
        else
            Ynew=Yn(1: length(Y));
            Yn=Ynew;
```

```
            error('Sampling rates of noise and Speech donot match');
        end
        %% HRTF convolve
        [ Sound_Front_Left]=conv(Y,h_aula_L);
        [ Sound_Front_Right]=conv(Y,h_aula_R);
        [ Noise_Left]=conv(Yn,h_aulanoise_L);
        [ Noise_Right]=conv(Yn,h_aulanoise_R);

        length(Sound_Front_Left);
        length(Noise_Left);

        xLeft=Sound_Front_Left + Noise_Left;
        xRight=Sound_Front_Right + Noise_Right  ;

        stimuli=[xLeft xRight];
        [max maxloc]=findmax(stimuli);  % find max absolute value and
location
        max;

        if (max <.001)
            stimuli=(1000*stimuli);
            ss=strcat('(max <.001) and -',Stimu);
        elseif (max <.01)
            stimuli=(100*stimuli);
            ss=strcat('(max <.01) and -',Stimu);

        elseif (max <.1)
            stimuli=(10*stimuli);
            ss=strcat('(max <.1) and -',Stimu);

        else
            stimuli=stimuli;
            ss=strcat('(max >.1) and -',Stimu);
         end


        if (max <1)
            stimuli=stimuli;
            yy=strcat('(max <1) and -',Stimu);
        elseif  (max < 1.5)
            stimuli=stimuli/1.5;
            yy=strcat('(max <1.5) and -',Stimu);
            %%%  disp('max <1.5')
        elseif   (max < 2)
            stimuli=(stimuli/2);
            yy=strcat('(max <2) and -',Stimu);
            %%% disp('max <2')
        elseif   (max < 2.5)
            stimuli=(stimuli/2.5);
            yy=strcat('(max <2.5) and -',Stimu);
            %%%  disp('max <2.5')
        elseif  ( max < 3)
            stimuli=(stimuli/3);
            yy=strcat('(max <3) and -',Stimu);
```

```
            %%%   disp('max <3')
        else disp('Outside the limit, Clipping of sound')
            yy=strcat('Outside the limit, Clipping of sound',Stimu);
        end


        xx= strcat(ss,'----',yy);

        fprintf(fwavefileProp,'%s\n\n', xx);
        Nbits=16;
        wavwrite(stimuli,fs,Nbits,Stimuliout);


        if mod(mk_folder,2)==0
            fprintf(fstimuli,'%s\n\n',strcat('lists: ' ,num2str(l),'-
' ,num2str(l+1),'/',Stimu));
        else
            fprintf(fstimuli,'%s\n\n',strcat('lists: ' ,num2str(l-1),'-
' ,num2str(l),'/',Stimu));
        end
    end

end
a=strcat(Stimuli_fol,'\','HRTFValue');
save(a,'air_aulanoise_L','air_aulanoise_R');
fclose(fspeech);
fclose(fstimuli);
fclose(fwavefileProp);
fclose('all');
```

2. Function to get the Noise stimuli for specific list –sentence for required SNR Ratio

```
function m_addnoise(speechfile, noisefile, nsnr,nf,outfile)
%nsnr is the Noise SNR
%nf : normalize to nf dB
%output: save the noise speech to output file

[x,Srate,nbits]=wavread(speechfile);

[n, Snrate,nnbits] = wavread(noisefile);


if  (Srate/Snrate)==1
    n =n;
else
    [P,Q]=rat(Srate/Snrate);
    n_new =resample(n,P,Q);
```

```
    n=n_new;
end


n_samples=length(x);
x=x*2^15;
% meen=mean(x);
% x= x - meen;

begin=randi([1 1001]);  %% ramdomisze begin %% changed from 600 to 1000
                                            %% since noise files is
                                            %% largethan 4 sec
%%n=(begin: begin + n_samples- 1);
n=n(begin: begin + n_samples- 1);
n=n*32768;

%----scale the noise file to get required SNR------------
se=norm(x,2)^2; %... signal energy
nsc=se/(10^(nsnr/10));


ne=norm(n,2)^2;  % noise energy

n=sqrt(nsc/ne)*n; % scale noise energy to get required SNR
ne=norm(n,2)^2;
fprintf('Estimated SNR=%f\n',10*log10(se/ne));

y= ( n)/ 2^15;       %% Since we only need Noise
wavwrite( y, Srate, nbits, outfile);
```

3. Function to Calculate the Binaural Room Impulse Response.

```
%%% this function calculates the HRTF for Different Room Sizes and Room
%%% Type as Specified

function [h_room_L,h_roomnoise_L,h_room_R, ...
    h_roomnoise_R,air_room_L,air_roomnoise_L,...
    air_room_R,air_roomnoise_R,aa,bb] = HRTF_room(azimuth_noise,...
    targetSrate,d,room,pathh)    % Azimuth angle (0° left, 90° frontal,
180° right)

airpar.fs = 48e3;
airpar.head = 1;  % With Dummy Head
airpar.rir_type = 1;
%                 '1': binaural (with/without dummy head)
%                          acoustical path: loudspeaker -> microphones
%                          next to the pinna
%airpar.room = 11;   % aula_carolina
%airpar.room = 5;   % Staircase
```

```
%airpar.room = 4;   % Lecture

airpar.room = room;
airpar.rir_no = d;
%airpar.rir_no = 3; %% Aula Carolina: {1m, 2m, 3m, 5m, 15m, 20m}
%%airpar.rir_no = 1;% Stairway: {1m, 2m, 3m}
%airpar.rir_no = 1;  % (5.56m) ->Lecture:  {2.25m, 4m, 5.56m, 7.1m,
%                         8.68m, 10.2m}

azimuthspeech=90;      %%% Direction of Speech
airpar.channel = 1;  %Left Ear

%%% direction of speech is infront therefore 90°
airpar.azimuth = 90; % Azimuth angle (0° left, 90° frontal, 180° right)
[h_room_L,air_room_L] = load_air(airpar);
airpar.azimuth=azimuth_noise;% Azimuth angle (0° left, 90° frontal, 180°
right)
[h_roomnoise_L,air_roomnoise_L] = load_air(airpar);

airpar.channel = 0;  %Right Ear
airpar.azimuth = 90; % Azimuth angle (0° left, 90° frontal, 180° right)
[h_room_R,air_room_R] = load_air(airpar);
airpar.azimuth=azimuth_noise;% Azimuth angle (0° left, 90° frontal, 180°
right)
[h_roomnoise_R,air_roomnoise_R] = load_air(airpar);

outputS={'h_room_L','h_roomnoise_L','h_room_R', ...
    'h_roomnoise_R'};

Fs=targetSrate;
fs=airpar.fs;

%%fs=airpar.fs;
if  (Fs/fs)==1
    %%do  nothing
else    %%Resample from  fs to Fs
    Y=h_room_L;
    [P,Q]=rat(Fs/fs,0.0001);
    Ynew =resample(Y,P,Q);
    Y=Ynew;
    h_room_L=Y;

    Y=h_roomnoise_L;
    [P,Q]=rat(Fs/fs,0.0001);
    Ynew =resample(Y,P,Q);
    Y=Ynew;
    h_roomnoise_L=Y;


    Y=h_room_R;
    [P,Q]=rat(Fs/fs,0.0001);
    Ynew =resample(Y,P,Q);
    Y=Ynew;
    h_room_R=Y;
```

```
    Y=h_roomnoise_R;
    [P,Q]=rat(Fs/fs,0.0001);
    Ynew =resample(Y,P,Q);
    Y=Ynew;
    h_roomnoise_R=Y;

    air_room_R.fs=Fs;
    air_roomnoise_R.fs=Fs;

end


Sp=strcat('Azimuth :-',num2str(azimuthspeech));
Np=strcat('Azimuth Noise :-',num2str(azimuth_noise));
S=strcat(Sp,Np);
fig1=figure();
subplot 211,plot(h_room_L)
subplot 212,plot(h_room_R)
xlabel(Sp);
fig2=figure();
subplot 211,plot(h_roomnoise_L)
subplot 212,plot(h_roomnoise_R)
xlabel(Np);
aa=strcat(pathh,'\',date,'-',num2str(azimuthspeech),'-speech','.png');
bb=strcat(pathh,'\',date,'-',num2str(azimuth_noise),'-noise','.png');
saveas(fig1,aa,'png');
saveas(fig2,bb,'png');

end
```

4. Add Noise to Speech only

```
%%Scale the noise to Required nsnr

%nf : normalize to nf dB
clc
clear all;
close all;
currentfolder = pwd;


%% First Run Trial_RoomIR.m then run this, Conv_HRTF_Audio_Script
disp('What list to start from ? ');
prompt = 'List No:';
result = input(prompt);
listNo = result;

prompt = 'What is the required Noise SNR ? ';
result = input(prompt);
```

```
nsnr = result;
currentlocation =pwd;
disp(' ')

inpath_speech =
'C:\Users\bhandary\Desktop\Testing\Database\IEEE\CleanVoice' ;
% location of clean files
inpath_noise = 'C:\Users\bhandary\Desktop\Testing\Database\noise' ;
% location of noise files
AirDataBase =  'C:\Users\bhandary\Desktop\Testing\Database\AIR_1_4' ;

output_folder= 'C:\Users\bhandary\Desktop\Testing\Reverb\AngieNoise-
TdB' ; %OutputFolder
if (exist( output_folder, 'dir')~= 7)
    if ~(mkdir( output_folder))
        error( 'Cannot create output directory');
    end
end

output_foldernoise=
'C:\Users\bhandary\Desktop\Testing\Reverb\AngieNoise-TdB-
noise' ; %OutputFolder
if (exist( output_foldernoise, 'dir')~= 7)
    if ~(mkdir( output_foldernoise))
        error( 'Cannot create output directory');
    end
end

fnoise= fopen( strcat(output_folder, '\', 'noiseConds.txt'), 'at');



noisefile = '2fsn-11062014-25k.wav';   %% Noise
fprintf(fnoise,'%s\n\n','----Noise +Speech Conditions ------' ,'noise
type :',noisefile);
% end
noisefile= strcat( inpath_noise, '\', noisefile);



%%  this one does per list of 10 sentences per list
mk_folder=1;
for l=listNo:1:listNo+1

    mk_folder= mk_folder+1;
        if mod(mk_folder,2)==0
        Stimuli_fol= strcat(output_folder, '\',num2str( l),'-
',num2str( l+1));
        mkdir(  Stimuli_fol);
        else
        end
    for i=1:1:10

        %outdir= strcat( output_folder, '\', num2str( i), ...
```

```
        %       '-', num2str( i+1));
        Speech=['S_' num2str(l) '_'  num2str(i) '.wav'];
        infile= strcat( inpath_speech, '\', Speech);
        noiseout=['n_' num2str(l) '_' num2str(i) '.wav'];
        outfile= strcat( Stimuli_fol, '\', noiseout);
        nf=64;  %nf : normalize to nf dB
        m_addspeech_noise(infile,noisefile, nsnr,nf,outfile);
        %begin=randi([1 251]);
        %addnoise(infile,noisefile, outfile,nsnr,begin,nf);
        fprintf(fnoise,'%s\n\n',strcat('lists: ' ,num2str(l),'-
' ,num2str(l+1),...
            ' : ',noiseout,' Snr: ' ,num2str( nsnr),'db/'));

    end
end
fclose(fnoise);
fclose('all');
```