PROJECT REPORT

FACTORS AFFECTING IMPAIRED COCKTAIL-PARTY LISTENING

IN NORMAL HEARING LISTENERS

AIISH Research Fund Project No. Ref: SH/CDN/ARF-29/2015-2016

Authors

Dr Chandni Jain

Mr Vikas M D

All India Institute of Speech and Hearing

Manasagangothri, Mysuru - 570006

March 2017

FACTORS AFFECTING IMPAIRED COCKTAIL-PARTY LISTENING IN NORMAL HEARING LISTENERS

Sanction No: SH/CDN/ARF-29/2015-2016, Dated 05.10.2015

Personnel	Ms. Amruta G Research Officer
Total Fund	Rs.4,33,000/-
Project duration	13 months
Principal Investigator	Dr Chandini Jain Lecturer in Audiology, Department of Audiology All India Institute of Speech and Hearing Mysore– 570 006, Karnataka, India E-mail: chandini_j_2002@yahoo.co.in
Co-Investigator:	Mr.Vikas M D Research Assistant Department of Audiology All India Institute of Speech and Hearing Mysore– 570 006, Karnataka, India E-mail: vikasmdslp@gmail.com

Acknowledgments

The investigators would like to acknowledge the Director, All India Institute of Speech and Hearing for being pivotal in granting AIISH Research fund for the project. We would like to acknowledge the HOD, Department of Audiology the resources from the department for testing. We also would like to extend our appreciation to the accounts section for providing support in maintaining the accounts. Our heartfelt gratitude also extends to all the participants involved in this study for their kind cooperation.

Contents

Abstract

Present study aimed to investigate various factors viz. age, suprathreshold processing and working memory affecting cocktail party listening in individuals with normal hearing sensitivity. A total of 92 participants with normal hearing sensitivity were included in the study. They were divided into two groups based on their age. 52 young normal hearing adults in the age range of 20 -40 years and 40 older normal hearing adults. Tests administered included speech perception in noise test to assess cocktail party listening, tests to assess suprathreshold processing namely gap detection thresholds, temporal modulation transfer function, inter-aural time difference, differential limen of frequency and ripple noise discrimination and working memory tests including digit span, digit sequencing and spatial selective attention. Results showed that older adults performed poorer than younger adults in all the tests. Also, temporal cues showed better relation with speech perception in noise compared to the spectral cues in older adults. This can be attributed to the disrupted neural synchrony which is due to poor frequency selectivity as observed through ripple noise discrimination. Individuals rely more on temporal cues due to poorer frequency resolution and phase locking mechanism and also on top down processes Results also showed that when there was a degraded speech input participants relied more on their higher cognition.

Chapter 1

Introduction

Humans have the ability to attend to the relevant information in the midst of other irrelevant messages despite of having multiple sound sources. This remarkable ability to attend to the relevant speech and simultaneously ignore the irrelevant messages or signals in the background is termed as *cocktail party effect* (Cherry, 1953). This is basically a selective attention ability which enables the listener to focus their attention to the desired target speaker, filter out and ignore the competing messages. However, Moray (1959) reported that not every individual demonstrate cocktail party effect, instead only 33% of people demonstrate this effect.

In a clinical set up, an individual is considered to be normal if he/she is having normal hearing threshold of audibility but our day to day communication does not rely on sound detection but instead it depends on the suprathreshold sound and feature extraction. It is reported that individuals have difficulty to understand speech or they face communication breakdown in situations like busy restaurants or in a cocktail party where there are multiple sound sources in the background despite of having clinically normal hearing thresholds. However, the reasons of large variability in speech understanding especially in background noise are not very clear for individuals with clinically normal hearing thresholds. Thus, it can be said that the audiogram or threshold of audibility at different frequencies is not a good measure to accurately predict the speech recognition performance in the presence of competing background despite of the age of the participant (Killon &Niquette, 2000; Souza et al., 2007).

There are several cues that would facilitate the cocktail party listening or selective attention. This includes the spatial separation between the target and the masker which can be due to the changes in the interaural time difference (Schneider et al., 2007), the prior knowledge about the location of the speaker (Kidd et al., 2005), familiarity of the target speaker's voice and topic (Helfer & Freyman, 2008), F_0 is another cue that differentiates one's voice from another and it becomes a critical cue for speech perception in noise. The inharmonicity serves as a cue to separate the sound sources (Culling & Darwin, 1993; Du et al., 2011) and the cognitive cue like working memory capacity also facilitates selective attention (Bregman, 1990).

Ruggles et. al. (2011) studied spatial selective attention on individuals with normal hearing sensitivity, which was correlated with temporal fine structure's ability. The subject's task was to understand a stream of speech in the presence of other streams coming from different directions. They concluded that the ability to encode temporal features at supra threshold levels influences the communication in adverse conditions in normal hearing individuals.

1.1. Factors affecting Speech Perception in Noise

1.1.1. The role of spectrotemporal processing on speech perception in noise. A complex broadband signal such as a complex tone or speech signal comprised of multiple frequencies are decomposed into a series of single narrowband frequencies by the auditory filters of our basilar membrane. The output of each filter has an envelope which has slow varying and temporal fine structure information. Both envelope and fine structure information are represented in the neural discharges but the temporal fine structure depends more on the phase locking of individual frequencies and it has an important contribution to understand speech in noise especially in fluctuating background noise (for review Moore, 2008). Moreover, it also has an eminent role in perceiving the pitch and frequency discrimination (Moore, 2003).

Speech perception in the presence of steady state background noise is difficult when compared to speech perception in fluctuating background noise which is known to be due to *dip listening*. In such situations when the background noise is fluctuating (non speech or speech), envelope cues may not be sufficient for speech intelligibility because envelope cues may not segregate the mixture of sounds when it is presented in different frequency bands wherein the temporal fine structure (TFS) provides information on changing F0 which helps in speaker identification (Zeng et al., 2005; Hopkins et al., 2008). When the auditory filters are broadened due to sensory neural hearing loss or due to aging, the output of these filters which extracts the temporal fine structure information will be varied which in turn degrades the central mechanism to decode this information. It is clear that the temporal fine structure information depends on the tonotopicity of basilar membrane and when this is affected the extraction of TFS information is also affected (Huss & Moore, 2005). Older adults may not be able to follow the rapid changes in an on-going speech stream which could be due to the inefficiency in their neural encoding of temporal events (Gordon – Salant, 2006).

Also, aging effects have been noted for those temporal events which requires binaural processing and the thresholds seems to be elevated for interaural time difference (ITD) and interaural phase difference (IPD) in them. Moreover, it was found that ITD thresholds for aged individuals did not vary significantly from that of the middle aged individuals which evidenced that temporal processing deficits appears in early age (Babkoff & Muchneik, 2002; Grose & Manro, 2010). Helfer and Vargo (2009) reported that there was a significant strong correlation between speech perception and gap detection thresholds and speech perception in noise was significantly different for younger and middle aged adults. This shows that the neural synchrony reduces with age and the temporal processing deficits can be seen in early age especially for binaural processing which in turn affects speech perception in noise (Grose & Manro, 2009).

Further, the ability to perceive speech is also related to the listener's ability to differentiate among frequencies (i.e., spectral resolution). The effects of reduced spectral resolution on speech perception can be inferred from the studies on cochlear-implant (CI) users and also on normal hearing listeners presented with vocoder-processed stimuli that simulate cochlear implant processing (Nelson et al., 2003). However, whether the reduced spectral resolution for individuals with normal hearing sensitivity results in adverse speech perception in noise is still not known.

1.1.2. Neural correlates of speech perception in noise. In a natural environment, humans are exposed to complex speech signal which is rich in active amplitude modulations and fast spectrotemporal fluctuations rather than simple pure tones. Our brainstem codes this complex features with precise temporal and spectral neural codes. Brainstem response to complex auditory stimuli such as speech (syllable) has two classes of separate time locked responses, transient and sustained which mimics the acoustic characteristics of the speech signal (Skoe & Kraus, 2010). It has shown to provide the information regarding the auditory processing of speech signal at the subcortical level, however, based on the constraints of the phase locking properties of the brainstem, the neural coding would be limited to the F_2 of the signal. The sustained portion of speech auditory brainstem response has shown to provide information regarding the neural response to the harmonic portion of the stimuli (periodic in nature) namely Frequency Following Response (FFR) which reflects the neural phase locking of the stimulus. Here, magnitude and timing measures become important wherein the timing measures would provide information on how the brainstem synchronously respond to the acoustic stimulus. In addition, the temporal measures such as the stimulus response correlations and quiet to noise correlations would provide information on how precisely the response mimics the stimulus and the effect of background noise on response waveform. The sustained peaks (D, E, and F) of FFR gives information about F_0 and the transient onset and peaks represents the filter characteristics which shows that there are separate neural streams to represent source and filter characteristics (Johnson, Nicol & Kraus, 2005).

Older adults frequently report of having difficulty in understanding speech especially in background noise despite of their normal hearing sensitivity (Jerger et.al, 1989). The speech ABR is an important measure to investigate the cocktail party problem in these individuals because human scalp recorded speech ABR represents the critical acoustic properties of the speech signal with considerable temporal and spectral precision (Du et al., 2011). Galbraith et al. (1995) demonstrated that when FFR is elicited for words, normal hearing listeners identified the stimuli with marked accuracy which indicated that it reflects the acoustic properties of the signal accurately. For identification of vowels, the main characteristic acoustic feature is its formants FFR encoding spectral peaks which are related to these formants (Krishnan, 2002). Not only vowels like sounds, it also represents the transient features, formant transition and sustained portion of a syllable (most commonly used /da/) (for review Johnson, Kraus & Nicol, 2005). It also encodes the pitch information for complex tones as well as for lexical tones (Greenberg et al., 1987, Krishnan et al., 2009).

Morever, in speech ABR, even when target speech is presented in multitalker babble or in the presence of noise even at low SNR, the neural response to target speech can be clearly differentiated from the masker. Russo et al. (2004) presented /da/ in quiet and in the presence of ipsilateral Gaussian white noise at +5 dB SNR in children and found that the transient peaks were most affected with delayed latencies and in most of the subjects it was absent, however, the sustained peaks were more resistant to the effects of background noise and remained stable. Among the sustained peaks, the F remained easily detectable in most of the subjects and latency was unchanged in background noise. However, the RMS amplitude, stimulus response correlation, F0 and F1 amplitudes were significantly reduced. Even though there was significant reduction in F0 amplitude it remained robust compared to other measures. The onset and offset peaks represent the transient, rapid time varying features of consonants which were most affected in background noise indicating a decrement in the neural synchrony to transient features. On the other hand, the FFR portion i.e. the sustained peaks which indicate the periodic vowels remained stable. The fourier analysis of the sustained portion gives the F0 amplitude which was robust even in background noise indicating that F0 is a major cue that helps us in identifying the speaker and the emotional tone in voice in background noise and it is more resistant to the detrimental effects of background noise. The FFR response to target and the masker can be easily separable as long as the F0 of the target and the masker remains the same.

Many times individuals also depend on their cognition in challenging listening situations to understand the target speech. This has also been measured using FFR, wherein, there are evidences that FFR can be modulated by selective attention to the target by marked changes in latency and its amplitude. Galbraith et al. (2003) reported that there was a significant change in the F0 amplitude to vowel stimulus when that target vowel was attended. Thus, it can be concluded that that some type of attention-related modulation is happening at the level of the brainstem.

FFR can also be used to study the effect of age on speech perception in noisy situations. Werff and Burns (2011) reported that there was a reduction in F0 and F1 amplitude in older adults indicating a significant deficit in phase locking of these components. Clinard and Tremblay (2010) showed a significant delay and reduced amplitude for the transient peaks with increasing age. In another study, Anderson et al. (2011) performed speech ABR for 170 msec /da/ stimuli in quiet and in the presence of multitalker babble at +10 dB SNR on older adults. Further, they grouped older adults into top performers and bottom performers based on their Hearing in Noise Test (HINT) scores. The authors reported that the top performers had significantly higher representation of F0 of the stimulus and it correlated with HINT scores. There was no difference in higher harmonics and the top scorers in HINT had similar FFR responses for quiet and noise condition. The authors inferred that pitch cues serve as an important cue for understanding the speech and the speaker in cocktail party listening.

Thus, FFR can be used as a useful tool to study speech perception in noise because of its unique features such as 1) FFR can replicate the stimulus with good temporal and spectral precision 2) FFR has good target specificity i.e. it can easily segregate the neural response to target from the masker 3) FFR can be used to study the cognitive influence on speech perception reflected as attention dependent modulations 4) FFR can be used to investigate the effect of age on speech perception in noise.(Du et al., 2011).

1.1.3. The role of cognition on speech perception in noise. With the advent of cognitive hearing sciences, the intersubject variability on speech perception in noise has also been explained by their cognitive ability. Fulton et al. (2015) reviewed the relationship between both peripheral hearing and central auditory process with cognition and provided a description of the possible mechanism underlying this relationship. Wide spread discrepancies exists regarding the relationship between peripheral hearing and cognition. Several correlational and cross sectional studies showed a relationship between the two (Lin et. al., 2011, 2013). However, there are other studies which did not show any relationship (Gates et. al., 1999, 2002). Studies which reported no relationship between peripheral hearing and cognition had small sample size, on the other hand those studies which reported to have correlation had large sample size (Fulton et al., 2015). Compared to the relationship between peripheral hearing and cognition; central auditory processing abilities showed significant relationship with cognition even in small samples.

As explained before, in cocktail party effect one needs to actively attend to the relevant messages, simultaneously filter out and ignore the irrelevant background noise. This ability to inhibit the distracting information is a fundamental aspect of our working memory capacity which is a crucial cognitive characteristic (Dempstar, 1991). Working memory is responsible for the manipulation and temporary storage of information which is necessary for complex cognitive tasks. According to Baddley (2000) working memory consists of four components namely central executive, phonological loop, visuospatial sketch pad and episodic buffer. Among these, central executive is an attention control system and the phonological loop is responsible for temporary storage and rehearsing of speech based information which is important for understanding speech in complex situations (Baddley & Hitch, 1974; Baddley, 2000).

Ronnberg, Rudner, Lunner and Zekveld (2010) designed a model known as Ease of Language Understanding (ELU) to address the issues related to speech perception in noise wherein they said that the input to the model is multisensory i.e. RAMBHO (Rapidly Automatically Multimodality Bound). When there are suitable conditions, the RAMBHO function mediates the implicit processing and helps in understanding the information by matching the input with the stored templates in long term memory. On the other hand, when there is suboptimum conditions which could be due to hearing loss or in noisy conditions, the input is distorted and RAMBHO fails to activate the stored representations in long term memory and a phonological mismatch occurs. When there is a mismatch between the speech input and phonological representation in the long term memory, an explicit processing and storage capacity is required to comprehend the meaning based on our previous knowledge and incomplete information. This mismatch is commonly seen in cocktail party situation, even for individuals with normal hearing. In such situations, individuals tend to depend more on their working memory capacity which becomes a good predictor for speech perception in noise. The ELU model tried to explain the interaction between implicit and explicit mode wherein implicit mode helps to pick up the speech units and explicit mode helps to reconstruct and fill the missing information (Ronnberg et al., 2010). Even though Rudner et al., (2008, 2009) explored the relationship between working memory capacity and speech perception in noise, the conclusions were drawn based on the findings in hearing impaired population wherein loss of audibility plays a role. Among 20 experimental studies examining the relationship between individual variability in speech perception in noise and their relation with cognition, only few studies showed a relation between the two adding to the variability and inconsistencies in the results (Akeroyd, 2008).

1.2. Need for the study

Based on the literature review, it can be noted that variability observed in speech perception difficulties in noise in normal hearing individuals could not be attributed to a single factor. Speech perception in noise depends on various factors such as age, temporal processing, and working memory. But there is a lack of literature related to the contribution of each factor in the speech perception difficulties in individuals with normal hearing sensitivity.

Studies have shown a link between cognition and speech perception in noise. However, in most cases cognition plays only a minor role and hearing loss plays the major role in speech perception in noise and also there is a large variation across studies which suggests for further examination. Studies have also shown a relation between TFS and speech perception in noise, however, the attempts are made only by altering the TFS in speech signal. There are few studies that have measured correlation between performance on TFS (measured using psycho-physical procedure) measures and speech perception in noise. Also, the role of spectral processing on speech perception in noise is still not explored. Moreover, the knowledge on the effect of age on all of these parameters is still limited and in the studies discussed above, any one or two possible factors have been studied to examine the effect on speech perception in noise. None of the studies investigated all the possible factors using both behavioral and electrophysiological measures. Hence, the present study tried to evaluate the role of each of these factors such as working memory, spectral resolution, temporal resolution and neural correlates to speech perception in fluctuating background noise and to examine which factor is more important in speech perception in noise.

1.3. Aim of the study

The aim of the current study was to investigate the effect of age, working memory, and supra-threshold processing (temporal and spectral processing) on the cocktail - party listening in individuals with normal hearing sensitivity

1.4. Objectives of the study

1. To assess the effect of age on cocktail-party listening.
2. To investigate the effect of suprathreshold processing on cocktail-party listening.
3. To investigate the effect of working memory on cocktail-party listening.

4. To investigate the neural mechanism behind cocktail party listening using FFR.

Chapter 2

Method

The current study investigated the effect of age, working memory and suprathreshold processing on cocktail party listening in individuals with normal hearing sensitivity. Moreover, the study also investigated which factor plays an important role in cocktail party listening.

2.1. Research Design

A cross-sectional descriptive research design was employed to achieve the aims. A between subject design was used to study the effect of age on speech perception in noise and the influencing factors. Within subjects design was used to assess which factor played the mojor role in cocktail party listening. Participants were selected using purposive convenient sampling technique and the purpose and nature of the study was explained to each participant. Informed consent was taken from all the participants and study adhered to the 'Ethical guidelines for bio-behavioural research involving human subjects' of All India Institute of Speech and Hearing (Basavraj & Venkatsan, 2009) and ethical committee approval was obtained prior to the commencement of the study.

2.2.Participants

A total number of 92 participants with normal hearing sensitivity participated in the study. They were divided into two groups based on their age. 52 young normal hearing (YNH) adults in the age range of 20-40 years with a mean age of 30 years (5.15) and 40 older normal hearing adults (ONH) in the age range of 60-80 with the mean age of 63 years (2.95) participated in the study. None of the participant had a significant history of otological disorders like middle ear infections, trauma to head and neck, no intake of otological drugs and significant medical history like renal, cardiac or metabolic diseases like diabetes mellitus. None of them had any complaints of reduced hearing or vestibular symptoms. Furthermore, participants of both the groups met the following criteria:

- Native speakers of Kannada.
- Bilateral A type tympanogram with ipsilateral and contralateral acoustic reflexes present atleast in 500 Hz, 1000 Hz and 2000 Hz to rule out middle ear dysfunction.
- Had formal education of atleast 10 years.
- Air conduction thresholds and bone conduction thresholds of lesser than 15 dBHL at all octave frequencies from 250 Hz to 8000 Hz for air conduction for YNH whereas for ONH air conduction thresholds lesser than 15 dBHL up to 2000 Hz and less than 30 dBHL at 4000 Hz and 8000 Hz (Kumar &Sangamantha, 2011; Snell, 1997)
- Speech recognition threshold within ±12 dB of pure tone average (four frequency Pure Tone Average 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz). The speech identification scores more than 80% assessed using Phonetically Balanced word list in Kannada for adults (Vijayalakshmi & Yathiraj, 2005) at 40 dBSL with reference to the SRT.

2.3.Instruments used

The following instruments were used for the study

- A caliberated diagnostic two channel diagnostic clinical audiometer (Maico, MA 53) with Telephonics TDH 39 headphone for obtaining air conduction thresholds, speech recognition thresholds and speech identification scores; Radio Ear B71 bone vibrator was used for obtaining bone conduction thresholds.

- A caliberated immittance audiometer (GSI – Tympstar version 2 middle ear analyzer) for assessing tympanometry and acoustic refelex thresholds.
- A Laptop (Sony Vaio SWE14123CNW model) installed with MATLAB version 7.10 (Mathworks Inc., 2010) for performing spectral and temporal tests.
- A Personal Computer (32 bit) installed with Biologic Navigator Pro (Natus Hearing Diagnostics) version 7.0 with software and attached hardware for electrophysiological measures.
- Sennheiser HDA 200 circumaural headphones with MX 141 adapter for presentation of all the test stimuli.
- Three Loudspeakers in horizontal array, Cubase SX, Aurora 8 and 16 A/D and D/A converter for presentation of stimuli for assessing spatial attention.
- Bruel and Kjaer 2270 Sound Level Meter, 6cc and 2cc couplers for caliberation.

2.4.Testing Environment

Basic audiological evaluation was carried out in a sound treated room (ANSI S3.1, 1991). All the experiment tests including speech perception in noise, working memory, tests to assess spectral, temporal processing and spatial selective attention were conducted in a quiet room.

2.5.Test materials

1. Temporal and spectral processing assessment. The maximum likelihood procedure (mlp) tool box and psychoacoustic tool box implemented Matlab v.7.10.0499 (R2010a) were used for assessing the spectral and temporal processing. Among which, the mlp tool box was used to assess gap detection test, modulation detection, difference limen for frequency and ripple noise discrimination while the psychoacoustic tool box was used for assessing interaural time difference.

2. Speech Perception in noise. The quick speech perception in noise test in Kannada (Methi, Avinash, & Kumar, 2009) was used for speech perception in noise assessment. The list 2 which has seven sentences at different SNR was used for the study.

3. Frequency following response. BioMRK protocol in Biologic Navigator Pro was used for recording Frequency Following Response. The Brainstem tool box was used to analyze the data and AEP – ASCII version 1.6.0 was used to obtain the data points.

5. Working Memory Assessment. Cognitive training module - Part 1 (Kumar & Sandeep, 2013) was used to assess the working memory through auditory digit span and sequencing.

2.6 . Stimuli and procedure

A written informed consent was obtained from all the participants for their willingness to participate in the study. A detailed case history was taken to rule out any otological, medical and family history of hearing loss. Questions related to history of middle ear infections, metabolic or systemic diseases, complaints of reduced hearing and vestibular complaints in different living situation and difficulty in speech in noise, memory, attention, intake of any ototoxic drugs, hypertension and cardiac problems etc. were asked. An otoscopic evaluation was done to rule out presence of impacted wax, to know the status of tympanic membrane and external auditory canal prior to the evaluations. A routine audiological evaluation was performed prior to the experimental tests to ensure normal hearing sensitivity. Further evaluation was done to assess factor which influence speech perception in noise. Five categories of tests were selected to fulfil this objective, which included,

a) Tests to assess suprathreshold processing
  - Spectral Processing
  - Temporal processing

b) Test to assess neural mechanism of speech perception

c) Tests to assess working memory

d) Test to assess cocktail party listening

Behavioral as well as electrophysiological measures were incorporated in the study. Stimuli for all the test were presented binaurally at 80 dBSPL. The following tests were used under each category

Tests to assess spectral processing	Difference limen for frequency (DLF) Ripple noise discrimination (RND)
Tests to assess temporal processing	Gap detection threshold (GDT) Modulation detection threshold (MDT) Interaural time difference (ITD)
Tests to assess working memory	Auditory digit span (forward and backward span) Auditory number sequencing (ascending and descending digit span) Spatial selective attention test

Test to assess neural mechanism of speech peception	Frequency Following Response (FFR) and FFT analysis of the same
Test to assess cocktail party listening	Quick speech perception in noise test

The maximum likelihood procedure implemented by a functional toolbox i.e. mlp toolbox (Grassi & Soranzo, 2009) in Matlab version 7.10 (Mathworks Inc., 2010) was used to determine the threshold for all the psychophysical tests. The maximum likelihood procedure employs a large number of psychometric functions and after each trial calculates the probability of obtaining the listener's response. The psychometric function yielding the highest probability is used to determine the stimulus to be presented at the next trial. Within about 12 trials mlp usually converges on a reasonably stable estimate of most likely psychometric function which can then be used to estimate the threshold (Green, 1990; 1993).

All the stimuli for psychophysical tests were generated at a sampling rate of 44100 Hz. A three interval alternate forced choice method was used to track the threshold on a 79.4% correct response criterion on psychometric function. Each trial had three blocks including two standard stimuli and a variable stimulus. The subject was expected to respond for the variable stimulus. All the stimuli for psychophysical tests were caliberated using Bruel and Kjaer 2270 using 6cc coupler at 80 dBSPL and presented through Sennheiser HDA 200 circumaural headphones. Before the actual testing, a proper instruction regarding the stimulus, response expected and 5-6 practice trials presented through speakers and appropriate feedback was given. A total of 30 trials excluding the practice trials were given to estimate the threshold.

The stimuli used for various tests are as mentioned below:

Table 2.1. *An overview of the tests carried out and stimuli used in the study*

Test	**Stimuli**
Difference limen for frequency	1000 Hz pure tone
Ripple noise discrimination	Gaussian noise
Interaural time difference	330 Hz pure tone
Modulation detection thresholds	Gaussian noise at 4, 8, 16, 32, 64 and 128 Hz
Auditory digit span and sequencing	English digits from 1-9 except 7
Spatial selective attention task	English digits from 1-9 except 7
Frequency following response	/da/ in quiet and in pink noise at +10 dB SNR
Quick speech perception in noise	Kannada sentences in the presence of four talker babble

The detail of each test is descibed below.

2.6.1. Tests to assess suprathreshold processing. Suprathreshold processing was assessed through spectral and temporal processing. Spectral processing was assessed through difference limen of frequency and ripple noise discrimination. Temporal processing was assessed through gap detection thresholds, interaural time difference and modulation detection thresholds.

Difference limen for frequency. The ability of the participant to detect the minimal difference between two stimuli that differ only in terms of its frequency was assessed. The DLF was obtained for 1000 Hz, 250 ms pure tone. The onset and offset of the tone was gated with 10 ms raised cosine ramps to avoid spectral splatter (Soranzo, Grassi & Massimo; 2014). The subjects were instructed that "you will be hearing three tones, among which one will have a slightly higher pitch than the other two. Your task is to identify which among the three tones is having the highest pitch". The starting level of presentation was 1100.1 Hz ($\Delta f = 100.1$ Hz). The first midpoint was 1000.1 Hz and last midpoint was 1100.1 Hz. The absolute difference limen for frequency (Δf) was noted.

Ripple noise discrimination. The frequency resolving power or spectral resolution was assessed using a complex stimuli i.e. ripples with peaks and dips in spectral domain which is a less time consuming procedure. The Gaussian noise of 500 ms is low pass filtered at 3000 Hz and sinusoidal ripples are generated by adding noise to itself at 5 ms delay. This delayed noise is attenuated by a variable amount. The standard noise is always 500 ms broad band noise with the same band pass filtering as the rippled samples but with uniform power spectrum. Both the standard and the delayed stimuli are equalized to average RMS power. The threshold is the attenuation in dB of the delayed noise (Soranzo, Grassi& Massimo; 2014). The subjects were instructed that "you will be hearing three noises among which one will be slightly different than the other two. Your task is to identify which is the odd noise among the three". The starting level was set to 30 dB with first midpoint at -20 and last midpoint of 0. The ripple discrimination threshold was noted.

Gap detection threshold. The smallest amount of silence that a participant can detect between two noise signals was assessed. A 750 ms Gaussian noise was used as the stimuli. The noise with a gap in its temporal center served as the variable stimulus and broad band noise of same duration with no gap/ continuous served as the standard stimulus. For the variable stimulus, the noise had 0.5 ms cosine ramps at the beginning and at the end of the gap to avoid spectral cues. The gap duration was varied according to the listener's performance. The subjects were instructed that "you will be hearing three noise stimuli wherein two will be continuous while one will have a gap or a period of silence.Your task is to identify which among the three is having the gap even if it is just detectable". The starting level of the gap was 64 ms with first midpoint at 0.1 ms and last midpoint at 64 ms. The gap detection threshold for all the participants were noted.

Interaural time difference. Speech recognition in noise is thought to be improved when the target speech and noise sources are separated in space. Hence spatial cues, like the interaural time differences are thought to be a major cue in improving the signal to noise ratio (Dubno, Ahlstrom Honwitz, 2002). Here, the interaural time difference was assessed for a 330 Hz pure tone of 250 ms. A two interval alternate forced choice was used where in both the tones have a certain ITD. The ITD of the variable tone (left tone) is varied. The same ITD value is used for the standard tone but with the opposite sign. The starting level was 300 wherein the first midpoint was at 0.0001 and last midpoint at 0.30. The subjects were instructed that "you will be hearing two tones one after the other. First you will be hearing in right following in left or vice versa. The task is to identify in which of your ears you heard the tone first or identify where the leading tone in is right or left". The threshold for ITD was noted.

Modulation detection thresholds. In this, the minimum amplitude modulation necessary to identify amplitude-modulated noise from an un-modulated white noise was assessed. The changes in the envelope of a speech signal is also thought to be an important cue for understanding speech especially in the presence of noise. A 1000 ms Gaussian noise was modulated at 4 Hz, 8Hz, 16 Hz, 32 Hz, 64 Hz and 128 Hz modulation frequencies. The MDT was measured in dB by using the following relationship:

$$\boxed{\text{Modulation detection thresholds in dB} = 20 \log10 m}$$

where, m= modulation index which ranges from 0- no modulation to 1- full modulation..

In three interval forced choice method, two blocks had standard un-modulated stimuli and one selected at random contained modulated stimuli. The participants had to identify which block had the modulated noise. The noise had 10 ms onset and offset cosine ramps and modulated and unmodulated noises were equated to total RMS power. The minimum and maximum amplitude modulation used was -5 dB and -35 dB. The temporal modulation transfer function (TMTF) was plotted for an average modulation detection thresholds for younger adults and older adults. The peak sensitivity and bandwidth was obtained for each individual using custom Matlab code.

2.6.2. Test to assess working memory. Working memory was assessed though auditory digit span, auditroy digit sequencing and spatial selective attention task.

Auditory digit sequencing and auditory digit span. This was done through ‘Auditory Cognitive Module’ (Kumar & Sandeep, 2013). Stimulus consisted of digits from one to nine except seven). The numbers were presented in random order with increasing level of difficulty with minimum of 2 digits and maximum of 10 digits with 250 msec of interstimulus interval. The test began by presenting 4 digits in random order and based on the performance of the subject the level of difficulty i.e. the number of digits increased or decreased. Two down one up method was used.

For auditory number sequencing the participants were presented with sequences of digits and they had to arrange the numbers heard in ascending (arrange numbers from lowest to highest) or in descending order (arrange the numbers from highest to lowest) and repeat them. The numbers of digits the participant can correctly recall in ascending and descending sequencing were noted. For e.g., if the participnat heard numbers ‘2984’ the expected response for ascending digit span is 2489 and expected response in descending digit span is 9842. If a number is repeated then the subject had to repeat it twice.

In auditory digit span test, the participants were presented with clusters of numbers and they are expected to repeat the numbers in the same order in forward digit span test and in the reverse order in the backward span test. For eg: if the participant heard ‘8943’, the expected response in forward digit span is 8943 and expected response in backward digit span is ‘3498’. Auditory working memory capacity is calculated as the total number of digits the subjects can recall in sequencing and digit span.

Spatial selective attention task. The selective attention was evaluated using digit streams from 1-9 except seven. Each stream consisted of eight digits. The stimuli were presented through loudspeakers arranged at 0 degree, +20 degree and -20 degree (horizontal) simultaneously. The subject was seated at the center facing the 0 degree azimuth speaker at 2 ft. distance. The stimuli was presented using Cubase SX music creation and production software installed in a personal computer and routed through Aurora 8 and 16 A/D and D/A converter and lynx mixer. The stimulus presented from each speaker was different. The output of each speaker was noted using Bruel and Kjaer 2270 sound level meter attached on the tripod stand from patient location i.e. at 2 feet distance from the speakers. The output of the speaker was adjusted such that output from all the speakers were at the same intensity level and the overall intensity when stimuli was presented from all the three speakers meet 80 dBSPL. The participant's task was to repeat the digits heard from the center speaker (at 0 degree azimuth) and ignore those heard from right and left speakers. The total number of correct responses was noted. For e.g. the particpant heard digit one from 0 degree speaker, digit eight from +20 degree and digit six from -20 degree speaker simultaneously. The subject's task is to repeat the digit one and ignore both 8 and 6.. Total of 16 targets were presented and the total number of correct responses from the participants were quantified out of 16.

2.6.3. **Test to assess neural mechanism of speech perception in noise.** In order to investigate the neural mechanism underlying the cocktail party effect in normal hearing adults, FFR was done for all the participants as it encodes the speech characteristics of signal, important for speech intelligibility (Du et al., 2011). FFR was done for 40 msec /da/ (Cunningham et al., 2001; Russo et al., 2004) stimulus in quiet and in pink noise at +10 dB SNR using BioMRK protocol in Biologic Navigator Pro version 7.0 (Natus Inc). The stimuli were presented binaurally using broad band insert. The acquisition parameters is mentioned in Table 2.2.

Table 2.2. *The acqusition papramenters used to elicit FFR*

Stimulus parameters	
Transducer	Insert earphones
Ear	Binaural
Polarity	Alternating
Intensity	80 dBSPL
Stimulus rate per second	10.1/s
Insert delay	0.80 msec

Stimulus	/da/ 40 msec (BioMAP-da.wav)
Masking type	Ipsilateral Quiet Pink noise (+10 dB SNR)
Recording Parameters	
Epoch	85.33 msec
Data points	1024
Pre/post	-15.00 msec
Blocking	0
Channel	1 channel
Gain	1,00,000
Artifact rejection	23.80 µV
Low filter	100 Hz
High filter	2000 Hz
Notch filter	Off
Montage	Input 1- Cz Input 2- right ear lobe Common ground – lower forehead

The absolute peak latency and amplitude were noted for each peak. Each response waveform was converted into ASCII file and FFT analysis was done using Brainstem Toolbox (Skoe & Kraus, 2010) in Matlab software version 7.10.0 (R2010a). The F0, F1 and F2 magnitude, the V/A slope, the SNR, the quiet to noise ratio, rms amplitude, stimulus to response correlations were noted for each response spectrum.

2.6.4. Test to assess cocktail party listening. Cocktail party listening was assessed through quick speech perception in noise test. Speech perception in noise was assessed for sentences in Kannada in the presence of four talker babble (Methi, Avinash and Kumar; 2009). The test consists of seven lists and each list consisted of seven sentenceswith five key words in each sentence. The signal to noise ratio was varied from +8 dB to -10 dB from first to seventh sentence in 3 dB steps. The subject was instructed to repeat the sentence as accurately as possible.The number of keywords correctly repeated was noted. Further, the SNR-50 was calculated, which is the SNR at which the subject can comprehend 50% of speech in noise. The SNR-50 was calculated using the Spearman Karber equation (Finney, 1952)

$$\text{SNR-50} = I + 1/2\ (d)\ (\#\text{correct}) / (w)$$

Where, I – initial presentation level = +8 dB

d- Attenuation step size = 3dB

w –number of keywords /decrement =5

c- Total number of correct key words (out of 35)

2.7. Statistical Analyses.

The data of the present study was subjected to statistical analyses using the Statistical Package for the Social Sciences (Version 17). Descriptive statistics was carried out to estimate the mean and standard deviation for all the tests. Following this Shapiro Wilk test of normality was done to analyze the normal distribution of the data in YNH and ONH participants. Since, many of the parameters were not normally distributed and in some parameters (e.g. DLF and RND), the standard deviation was more than half of the mean; non parametric tests were administered. The data was tested for outliers but no significant outliers were identified. Mann Whitney U test was administered to test the significance between two independent groups (YNH and ONH) on different dependent variables. Spearman correlation was used to explore the relationship between Quick SIN and each of the tests within each group. A mixed design was used wherein the correlation between each tests under spectral, temporal and working memory categories and Quick SIN were the within subject variables and the effect of age was considered as the between subject (YNH vs. ONH) variable.

Chapter 3

Results

The present study investigated the effect of age on suprathreshold processing and cognition and factors which contributed in impairment in cocktail party listening. The results of the study are discussed under the following headings:

1) Effect of age on speech perception in noise (SPIN)
2) Effect of suprathreshold processing on SPIN
3) Effect of working memory on SPIN
4) Neural basis for SPIN

3.1. Effect of age on Speech Perception in Noise.

The number of keywords correctly identified by each participant and the SNR 50 was calculated using Spearman Karber equation was noted and the group data was analysed. Figure 3.1 shows the mean number of keywords correctly identified by both the groups along with standard deviation (SD) and Figure 3.2 shows the SNR-50 values for both the groups. It is evident from the Figure 3.1 and 3.2 that the YNH have better speech perception in noise performance compared to ONH. The SNR-50 in YNH adults ranged from -10.30 to -3.70 dB SNR with a median of -6.40 dB and in ONH adults ranged from -6.10 to + 2.30 dB SNR with a median of -2.50 dB. This shows that the ONH adults require higher signal to noise ratio compared to young normal hearing adults to understand speech in the presence of fluctuating background noise i.e. in situations like cocktail party.

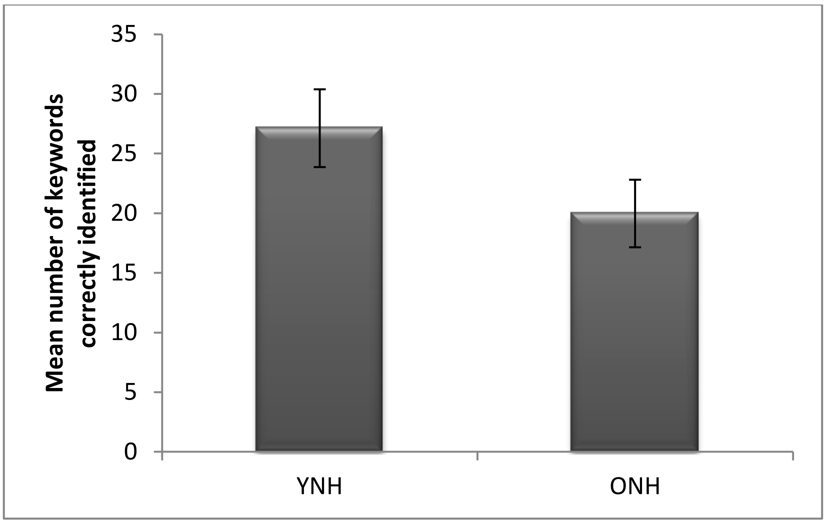


Figure 3.1. The mean and one standard deviation (SD) error bar for the mean number of keywords in quick speech perception in noise test by ONH and YNH.

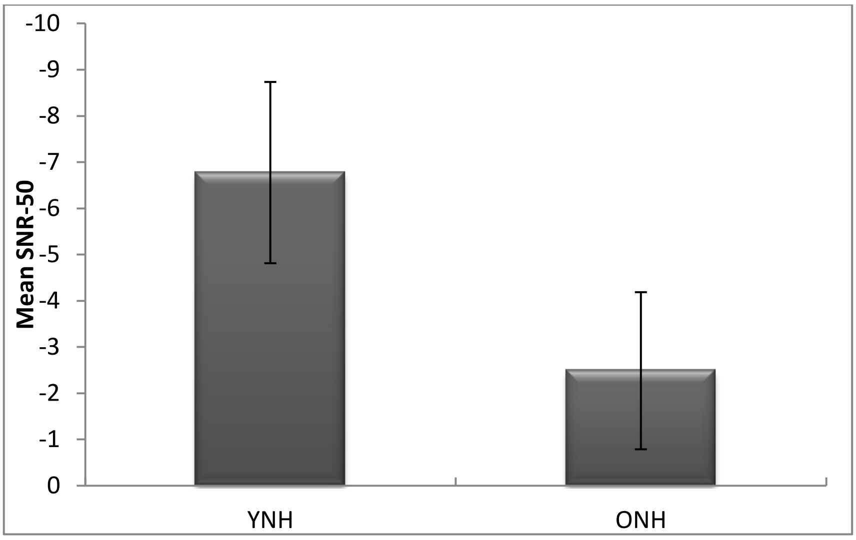


Figure 3.2. The mean and one SD error bar for SNR-50 in YNH and ONH.

Further Mann Whitney U test was administered to assess the difference in quick speech perception in noise scores between the groups (YNH and ONH). The results revealed that there was a significant difference in speech perception scores between both the groups (Z=-7.508, p<0.01). The results revealed that the ONH adults performed significantly poorer than the YNH adults in quick speech perception in noise test and required higher SNR to comprehend the sentences in a natural cocktail party listening condition.

3.2. Effect of Suprathreshold processing on Speech Perception in Noise

3.2.1. Role of temporal processing on speech perception in noise. The temporal processing was assessed using gap detection test (GDT), interaural time difference (ITD) and modulation detection threshold (MDT) for different modulation frequencies. Further, the peak sensitivity and bandwidth was calculated from the modulation detection thresholds and the temporal modulation transfer function (TMTF) was plotted for both the groups. Figure 3.3 represents the mean MDT across various frequencies in YNH and ONH adults respectively. The mean and one SD of GDT, ITD, peak sensitivity and bandwidth for YNH and ONH are shown in Figure 3.4. It can be inferred from the Figure 3.4 that the ONH adults had elevated thresholds for all the temporal measures compared to YNH adults indicating an age related decline in temporal processing. Further Mann Whitney U test showed that there was a significant effect of age on GDT (Z=-6.785, p<0.01), ITD (Z = -3.981; p < 0.01), peak sensitivity (Z=-6.748, p < 0.01) and band width (Z=-2.688, p< 0.01).

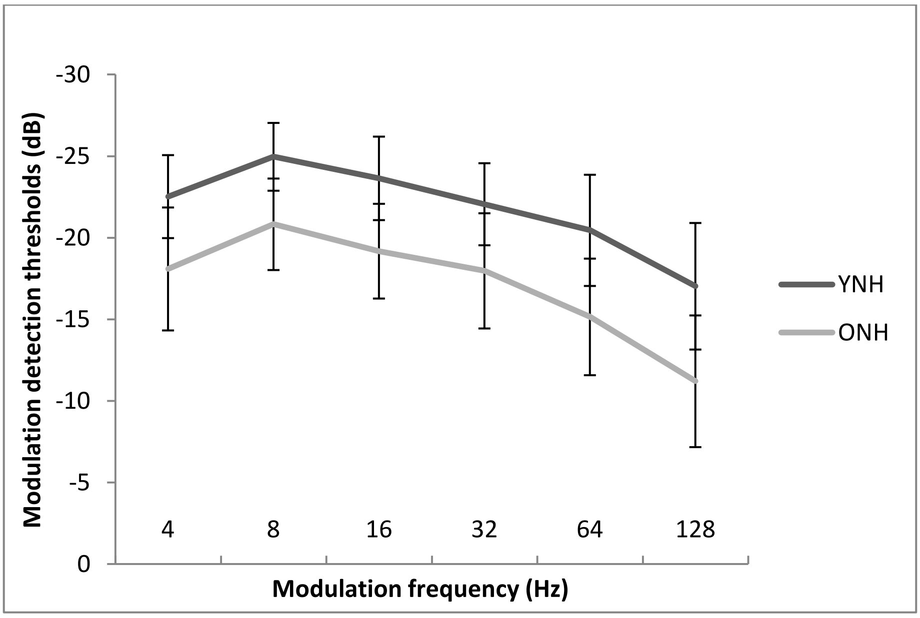


Figure 3.3. The mean and one SD error bar for MDT across frequencies in YNH and ONH.

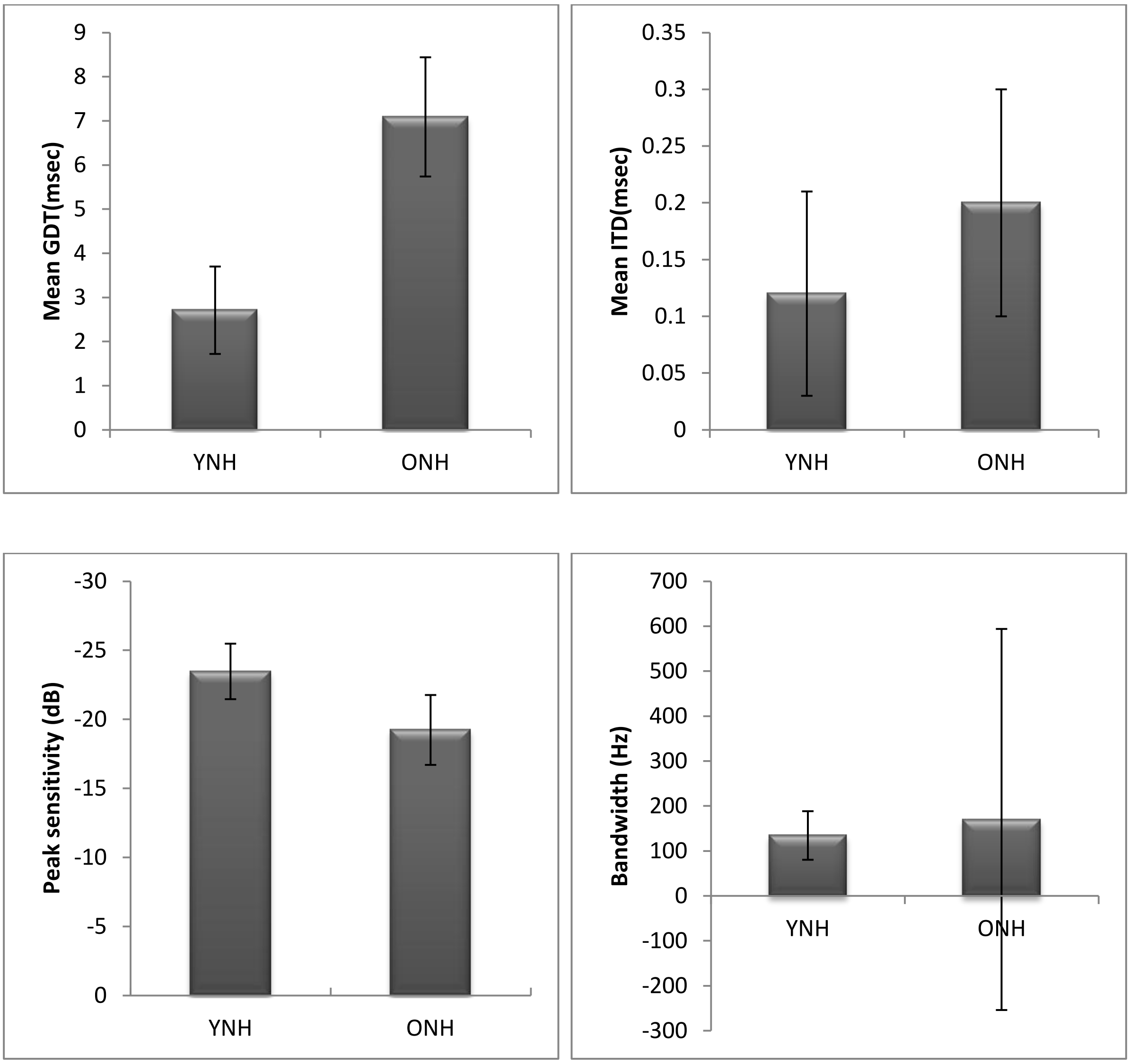


Figure 3.4. The mean and one SD error bar for GDT, ITD, peak sensitivity and band width in YNH and ONH.

The relationship between the various temporal processing skills and the speech perception in noise was also assessed. Each test that taps the temporal processing skills (GDT, ITD and MDT) were compared to the scores of quick speech perception in noise scores. Spearmann correlations were used to determine the relationships between speech perception in noise and temporal processing ability. Table 3.1 shows the correlation coefficients between speech perception in noise and different temporal processing tests. The results revealed that there was a significant, strong negative correlation between GDT and ITD with SNR-50 in YNH and ONH adults. The scatter plot of the significant correlation is plotted in Figure 3.5 and 3.6 for YNH and ONH. To conclude, GDT had maximum correlation with SPIN in all temporal proceesing abilities in both YNH and ONH.

Table 3.1.

Result of Spearman rank correlation between speech perception in noise and temporal tests

Test	GDT	ITD	PS	BW
YNH	-.74**	-.69**	-.22	-.26
ONH	-.67**	-.55**	.158	.23

Note:** Correlation is significant at the 0.01 level (2-tailed).

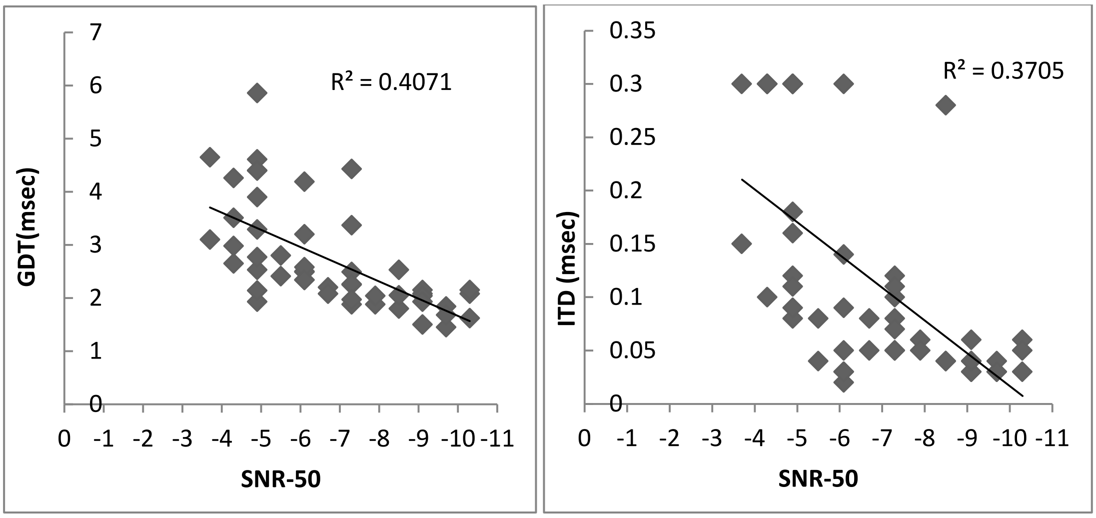


Figure 3.5. Scatter plot representing the relationship between GDT and ITD with SNR-50 in YNH adults.

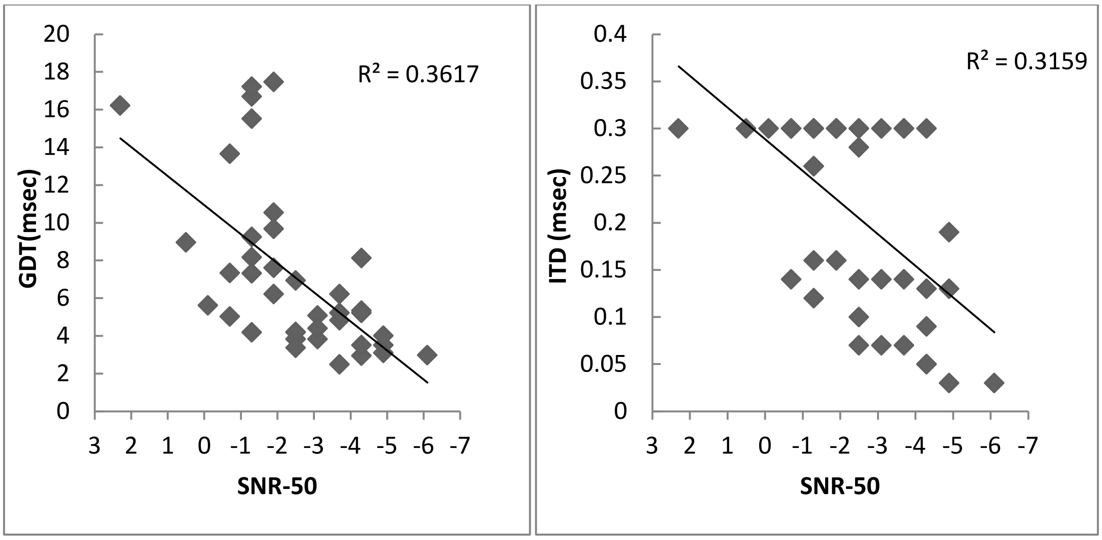


Figure 3.6. Scatter plot representing the relationship between GDT and ITD with SNR-50 in ONH adults.

3.2.2. Role of spectral processing on speech perception in noise. 3.1.3. Spectral processing was assessed using differential limen of frequency (DLF) and ripple noise discrimination (RND). The mean and one SD of DLF and RND in YNH and ONH is shown in Figure 3.7. It is evident from the figure that Δf is higher for ONH compared to YNH and ripple noise discrimination abilities was poorer in ONH compared to the YNH. Further, the Mann Whitney U test revealed that there was an significant effect of age on DLF (Z=-5.679, p<0.01) and RND (Z= -3.240, p<0.01).

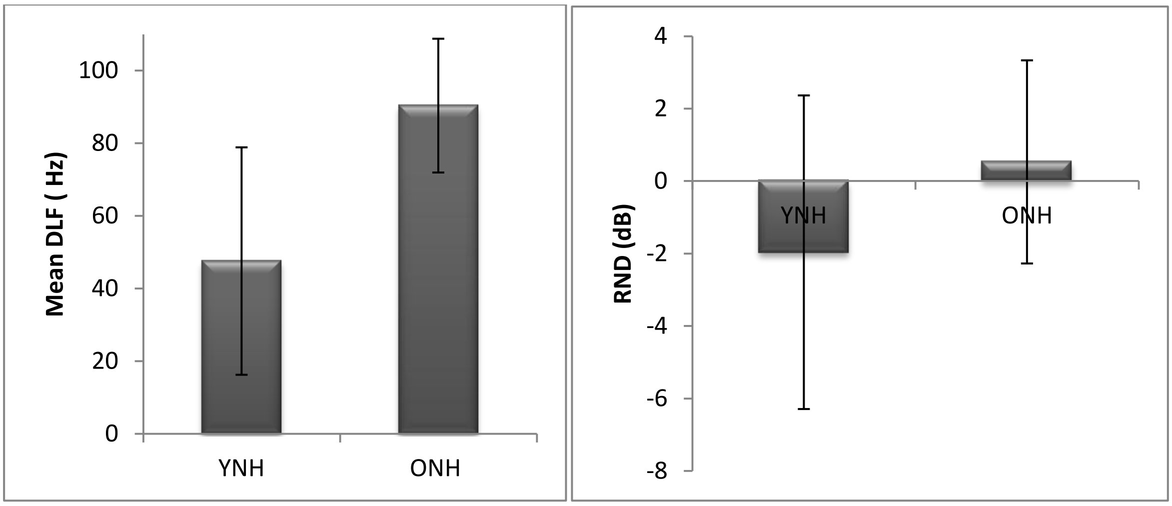


Figure 3.7. The mean and one SD of absolute DLF and RND in YNH and ONH adults

The results of Spearmann correlations revealed that there was a significant moderate, negative correlation between SPIN and tests to assess spectral resolution (DLF and RND) in YNH adults. However, in ONH adults DLF and RND showed very weak and weak correlations with SPIN. Table 3.2 shows the correlation coefficients between speech perception in noise and different spectral processing tests. The scatter plot of the significant correlation is plotted in Figure 3.8 and 3.9 for YNH and ONH.

Table 3.2.

Result of Spearman rank correlation between speech perception in noise and spectral tests (DLF &RND)

Test	DLF	RND
YNH	-.50 **	-.42 **
ONH	-.09	-.37*

Note:** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

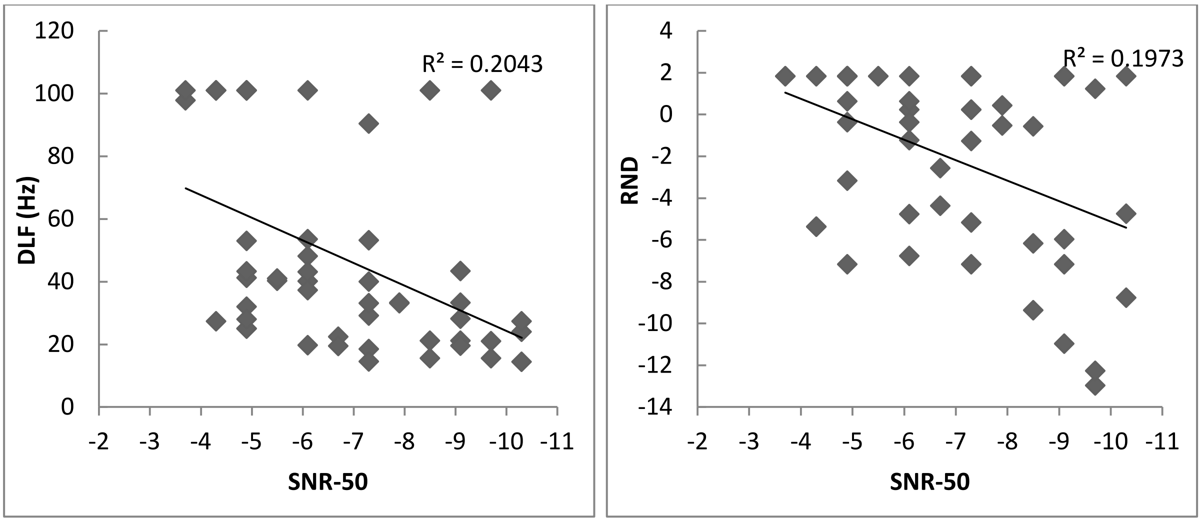


Figure 3.8. Scatter plot representing the relationship between DLF and RND with SNR-50 in YNH adults.

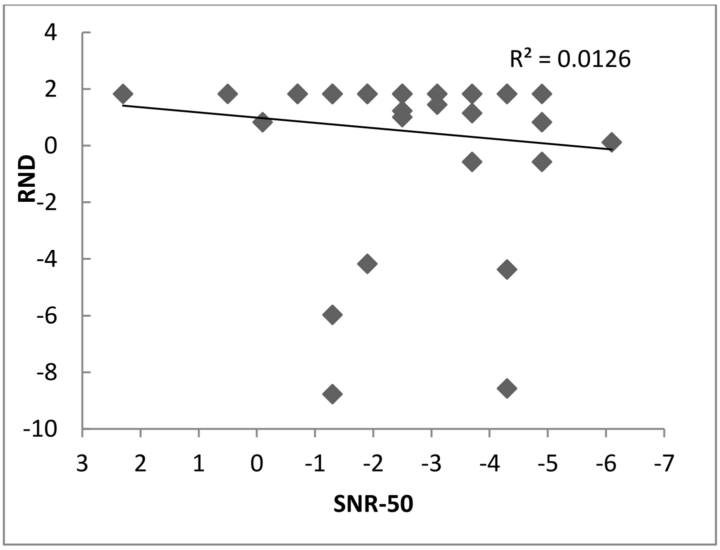


Figure 3.9. Scatter plot representing the relationship of RND with SNR-50 in ONH adults.

3.3. Role of Working Memory capacity on Speech Perception in Noise

The working memory capacity of the individuals was assessed using auditory digit span, auditory number sequencing tests and spatial selective attention task (SSA). The mean and one SD of forward, backward, ascending, descending digit span and SSA task is shown in Figure 3.10 for both the groups. From 3.10 it is evident that YNH had better working memory scores for all the test compared to the ONH. The results of Mann Whitney U test revealed that there was a significant difference between both the groups in forward digit span (Z= -4.464, p< 0.01), backward digit span (Z=-4.358, p<0.01,), ascending digit span (Z=-2.249, p<0.01), descending digit span (Z=-4.130, p<0.05) and SSA (Z= -3.305, p<0.01).

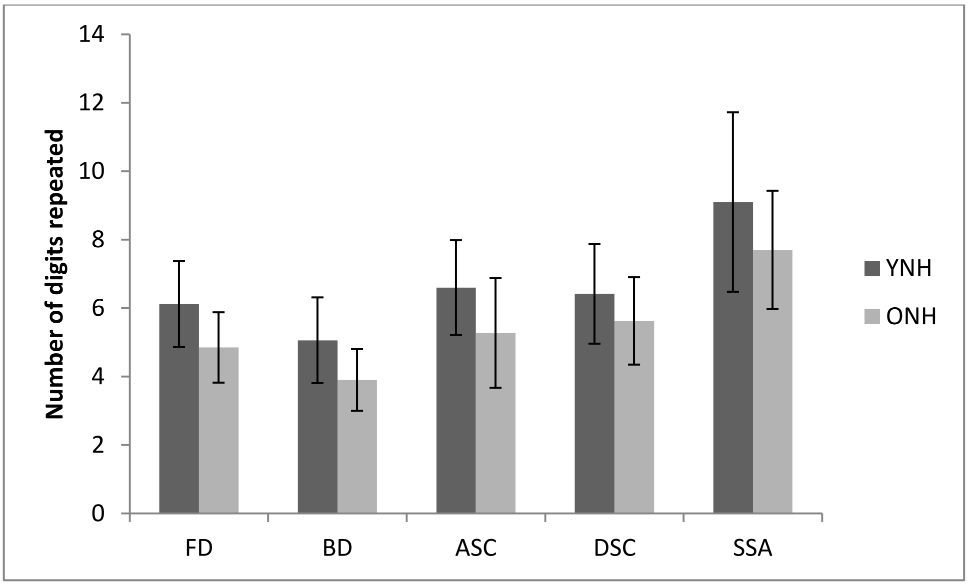


Figure 3.10. The mean and SD of working memory tests in YNH and ONH adults (FD- forward digit test, BD- backward digit test, ASC- asceding digit test, DSC- descendign digit test, SSA- spatial selective attention task)

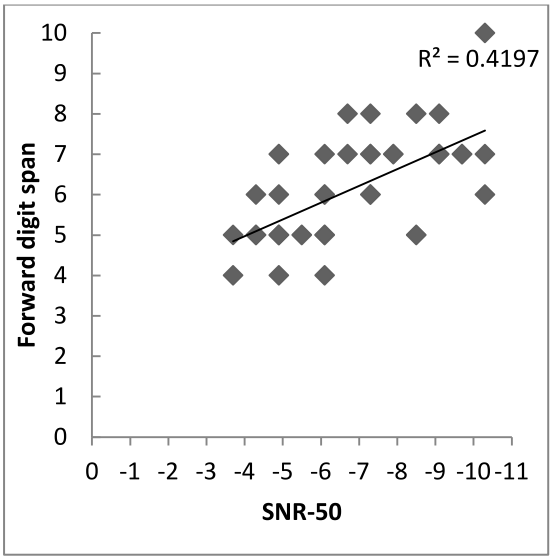
Spearmann correlation between working memory tests and SPIN was done in YNH and ONH. Table 3.3 shows the correlation coefficients between speech perception in noise and different working memory tests. The scatter plot of the significant correlation is plotted in Figure 3.11 and 3.12 for YNH and ONH. Results showed that YNH adults had a significant, strong, positive correlation with forward digit span and descending auditory digit span test and a moderate correlation with backward digit span and ascending number sequencing In addition, spatial selective attention showed no significant correlation with SPIN. These indicates that working memory capacity or cognition has a significant role in speech perception in noise for YNH. In contrast, ONH adults did not show relation between forward, backward digit span and spatial selective attention test with SPIN. However, ONH showed a significant, positive correlation with ascending and descending digit span tests. The results in ONH adults shows that a working memory test that requires higher cognition will have more correlation with speech perception in noise.

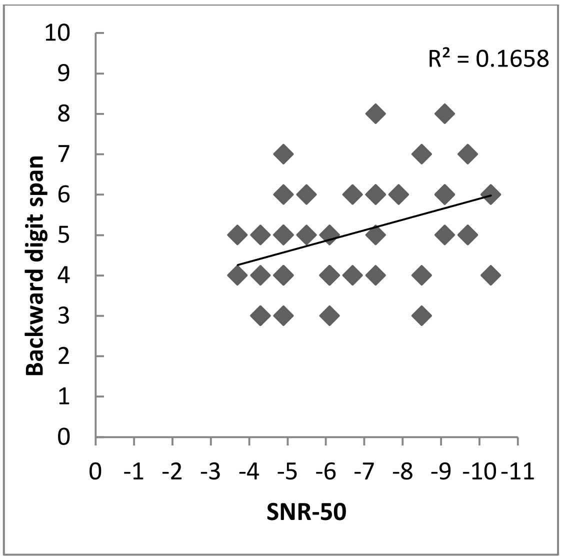
Table 3.3.

Result of Spearman rank correlation between speech perception in noise and working memory tests

Test	FD	BD	ASC	DSC	SSA
YNH	.65 **	-.44 **	.59**	.79**	.03
ONH	.12	.13	-.33**	.63**	-.009

Note:** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).





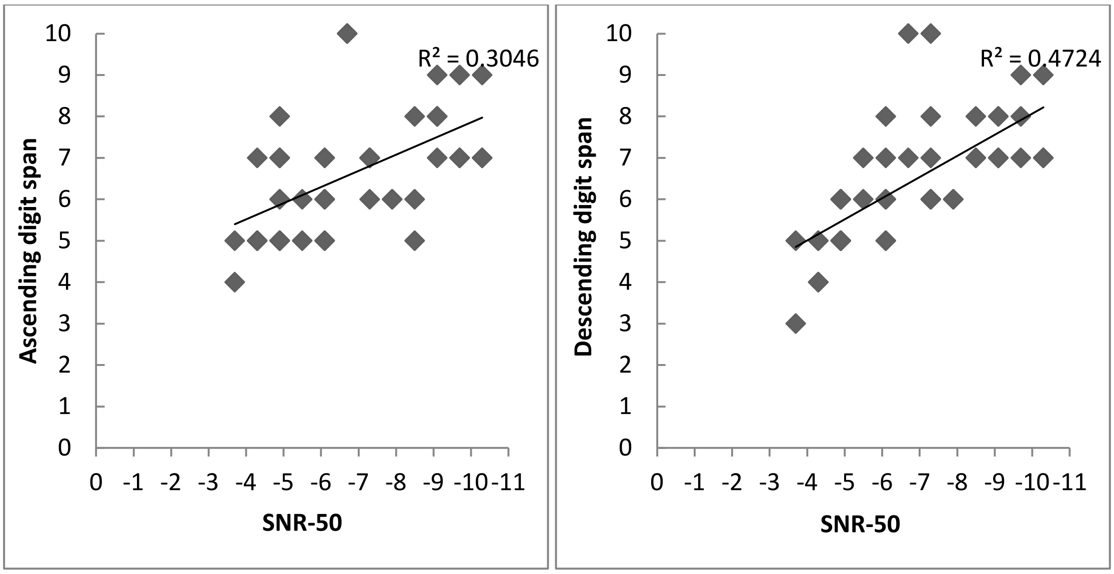


Figure 3.11. Scatter plot representing the relationship of working memory tests (forward, backward, ascending and descending digit span) with SNR-50 in YNH adults

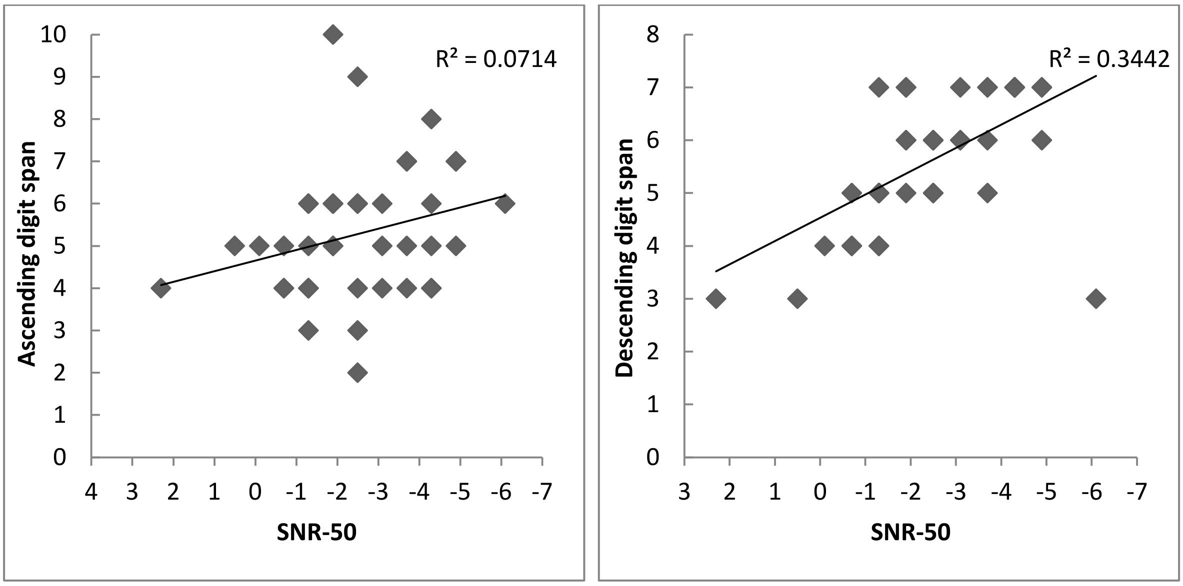


Figure 3.12. Scatter plot representing the relationship of working memory tests (ascending and descending digit span) with SNR-50 in ONH adults

3.4. Neural basis for Speech Perception in Noise

The neural mechanism behind speech erception in noise was assessed through FFR. The FFR was recorded for 40 msec /da/ stimuli at 80 dBSPL presented binaurally in alternating polarity with 4000 sweeps. The transient and sustained portions of FFR response waveforms were analysed. The data was analysed for the effect of aging, effect of background noise and relationship between the brainstem encoding and behavioural speech perception in noise in younger and older adults. Among the transient portions, the latency and amplitude of V-A complex was considered as the main parameter which reflects the neural encoding of the burst portion of the stop consonant. FFT analysis of the sustained responses in the region of 10 to 40 msec was done by MATLAB (version 2010a, Mathworks Inc; Natick, MA) routines developed by Erika Skoe and Trent Nicol at North western University (Brainstem Toolbox, 2008). In the offline analysis of the sustained portion of FFR or speech ABR, the following parameters were noted (Russo et al., 2004):

1) RMS amplitude- it gives the magnitude of neural activation over a given time period. The RMS amplitude of the response was divided by the pre-stimulus RMS, the quotient of which gives the SNR.

2) Stimulus response correlation which compares the overall morphology and timing of stimulus and response. It shows how the response waveform mimics the stimulus waveform.

3) Fourier analysis between 10 to 40 ms epoch assess the amount of neural activity in these frequency regions which is indicated by F0 amplitude and F1 amplitude.

4) Quiet to noise response provides a way to quantify the effect of background noise on sustained response in terms of morphology and timing.

Table 3.4 shows the amplitudes and latencies of discrete peaks- waves V, A and the F0, F1 amplitude, SNR and rms values in quiet and noise for both the groups. Table 3.5 shows the results of Mann Whitney U test to assess the significance between the two groups for various amplitude, latency and FFT paramterers of FFR. It can be seen from the table that bothe the groups differed significantly in FFR in quiet conditions and for few parameters in noise conditions in both the groups. Further Wilcoxon signed rank test was used to compare the FFR response in quiet vs. noise condition and results showed that there was a significant difference ($p < 0.001$) between the transient and sustained portions in quiet and in the presence of noise in both the groups as shown in Table. 3.6

Table 3.4.

The Mean and SD of various latency, amplitude and FFT parameters of brainstem responses of speech in YNH and ONH adults

Parameters		YNH		ONH	
		Mean	SD	Mean	SD
Latency	Wave V (quiet)	6.25	0.24	6.59	0.26
	Wave A (quiet)	7.38	0.31	7.78	0.39
	Wave V (noise)	6.87	0.45	7.23	0.81
	Wave A (noise)	8.13	0.37	8.48	0.78
Amplitude	Wave V (quiet)	0.12	0.07	0.07	0.03
	Wave A (quiet)	-0.14	0.06	-0.09	0.04
	Wave V (noise)	0.07	0.05	0.04	0.02
	Wave A (noise)	-0.08	0.05	-0.06	0.03
FFT	F0 amplitude in quiet	3.40	1.07	2.69	1.17
	F1 amplitude in quiet	0.32	0.15	0.28	0.10
	SNR in quiet	3.93	1.82	2.22	0.69
	rms in quiet	0.06	0.02	0.04	0.01
	F0 amplitude in noise	2.13	0.93	1.91	1.13
	F1 amplitude in noise	0.21	0.08	0.24	0.16
	SNR in noise	2.24	1.17	1.37	0.53
	rms in noise	0.04	0.02	0.04	0.02

Table 3.5.

The Z value and significance level for latency, amplitude and FFT parameters of brainstem responses of speech between YNH and ONH adults

Parameters		Z-value	Significance level
Latency	Wave V (quiet)	-5.513	**.000**
	Wave A (quiet)	-4.439	**.000**
	Wave V (noise)	-2.143	**.032**
	Wave A (noise)	-2.458	**.014**
Amplitude	Wave V (quiet)	-4.915	**.000**
	Wave A (quiet)	-4.787	**.000**
	Wave V (noise)	-3.359	**.001**
	Wave A (noise)	-1.702	.089
FFT	F0 amplitude in quiet	-2.797	**.005**
	F1 amplitude in quiet	-1.374	.169
	SNR in quiet	-5.190	**.000**
	rms in quiet	-4.867	**000**
	F0 amplitude in noise	-1.888	.059
	F1 amplitude in noise	-.674	.500
	SNR in noise	-3.524	**.000**
	rms in noise	-1.574	.115

Table 3.6.

The Z value and significance level for latency, amplitude and FFT parameters of brainstem responses between quiet and noise in YNH and ONH adults

Parameters (quiet vs noise)		YNH		ONH	
		Z-value	Significance level	Z-value	Significance level
Latency	Wave V	-5.907	.000	-5.162	.000
	Wave A	-6.017	.000	-5.304	.000
Amplitude	Wave V	-5.007	.000	-4.485	.000
	Wave A	-5.269	.000	-3.974	.000
FFT	F0 amplitude	-6.031	.000	-5.303	.000
	F1 amplitude	-5.334	.000	-4.195	.000
	SNR	-6.031	.000	-5.288	.000
	rms	-5.652	.000	-3.780	.000

Spearmann correlation to assess the correlation between various parameters of FFR with SPIN showed that the SPIN correlated with f0 and f1 parameters in both quiet and noise for YNH and ONH. Table 3.7 shows the correlation coefficients between speech perception in noise and different parameters of FFR. The scatter plot of the significant correlation is plotted in Figure 3.13 and 3.14 for YNH and ONH.

Table 3.7.

Result of Spearman rank correlation between speech perception in noise and Various parameters of FFR in YNH and ONH

Parameters		YNH	ONH
Latency	Wave V (quiet)	-.186	**-.434***
	Wave A (quiet)	-.114	-.238
	Wave V (noise)	-.117	-.210
	Wave A (noise)	-.224	-.128
Amplitude	Wave V (quiet)	.312	.060
	Wave A (quiet)	-.233	-.209
	Wave V (noise)	.119	.140
	Wave A (noise)	-.384	-0.091
FFT	F0 amplitude in quiet	**.641****	**0.582****
	F1 amplitude in quiet	-0.016	.281
	SNR in quiet	.196	.228
	rms in quiet	**.462****	**.399***
	F0 amplitude in noise	**.507****	**.479****
	F1 amplitude in noise	-.126	.200
	SNR in noise	.193	.201
	rms in noise	**.466****	.207

Note:** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

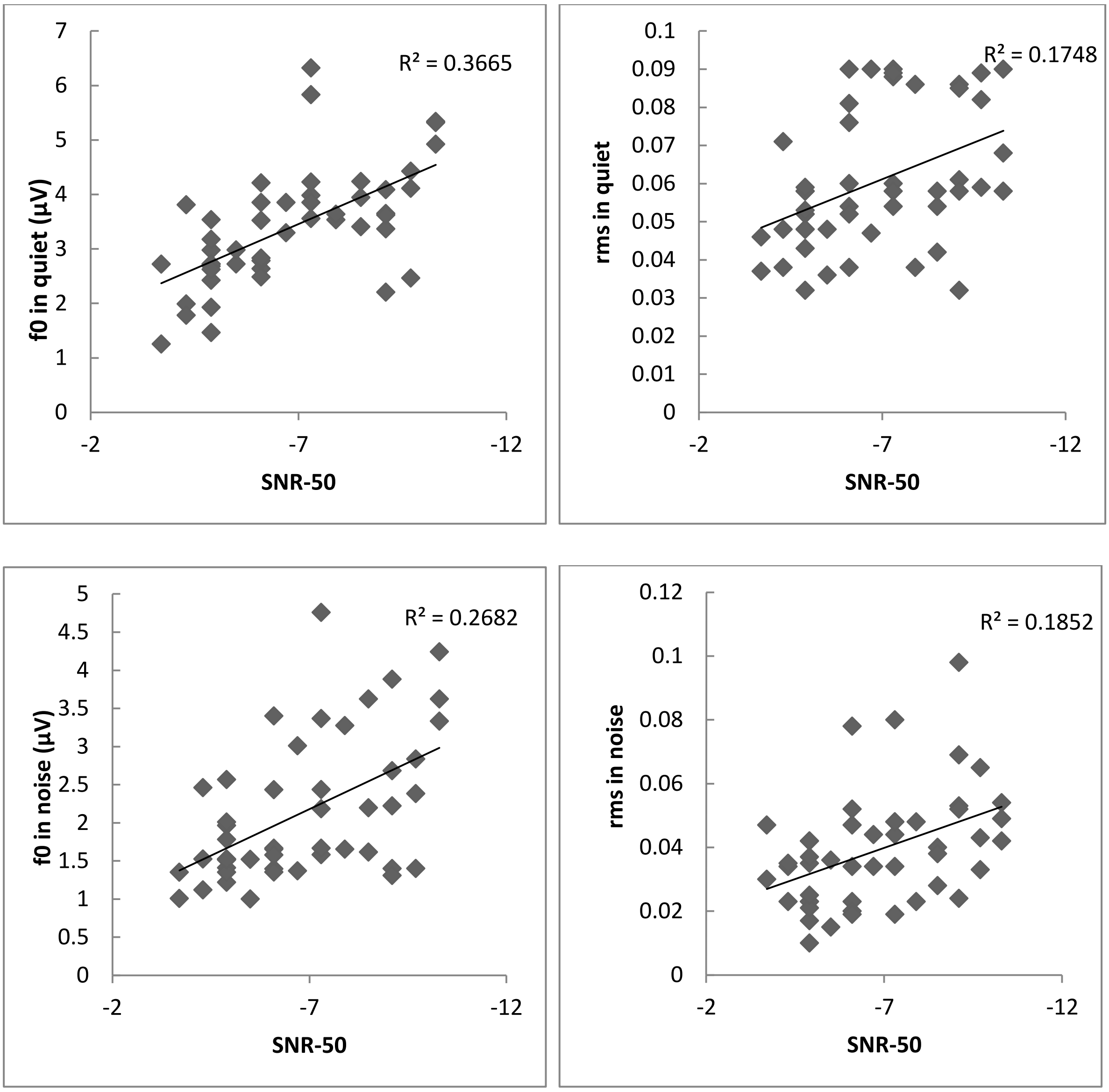


Figure 3.13. Scatter plot representing the relationship of f0 and rms in quiet and noise with SNR-50 in YNH adults.

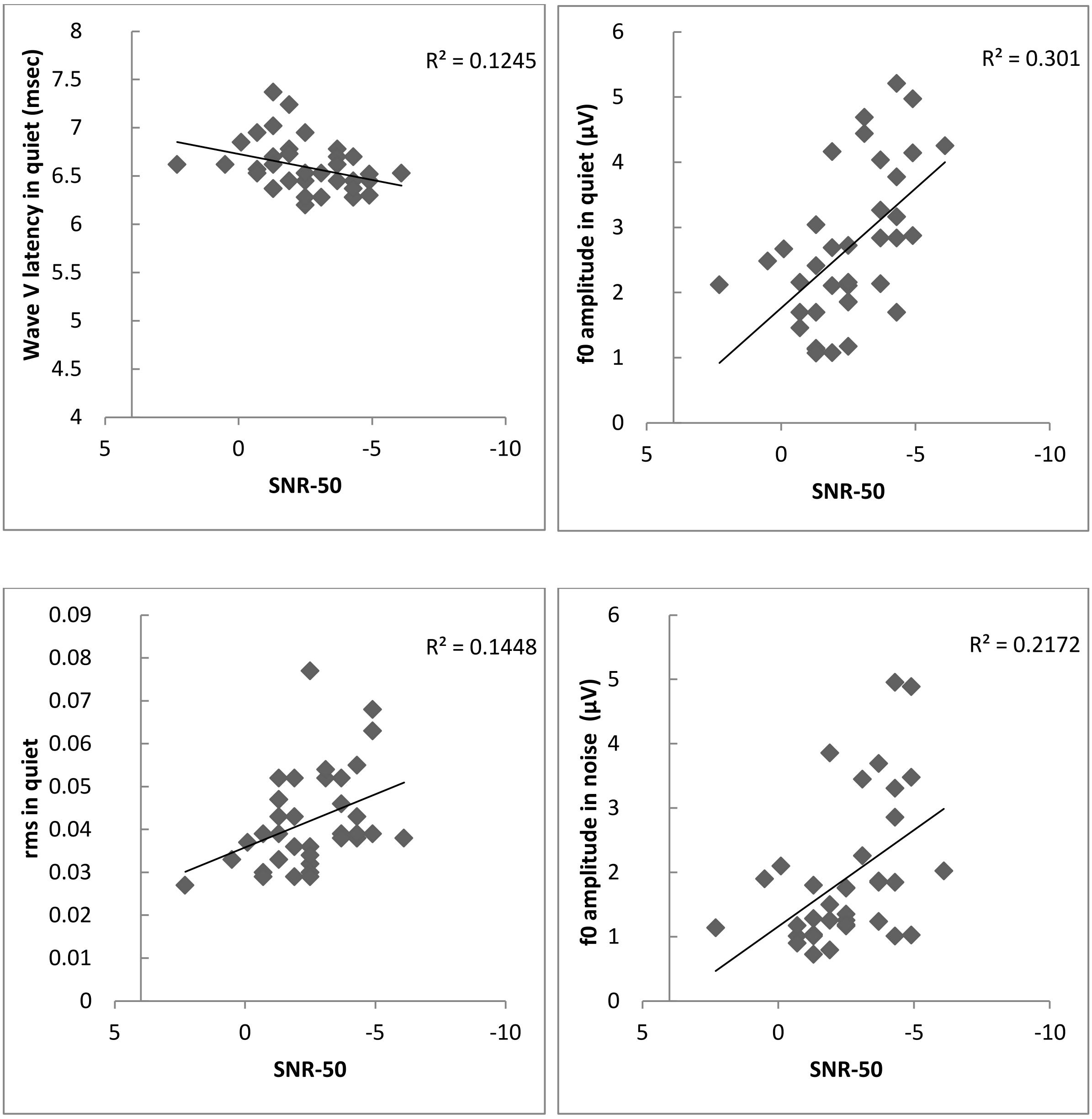


Figure 3.14. Scatter plot representing the relationship of wave V latency, f0 and rms in quiet, f0 in noise with SNR-50 in ONH adults.

Based on the morphology, amplitude and timing of the FFR waveforms and its correlation with quick speech perception in noise test wherein those who scored high in quick speech perception in noise had better waveform morphology with clearly distinguishable peaks with good amplitude. Hence, for each group, individuals were divided into top and bottom performers based on their SPIN scores. Those individuals who scored greater than the median value were considered as top performers and those who scored below the median fell into bottom performer group. The top and bottom performers showed significant difference in their SPIN scores in both younger and older adults ($p< 0.001$). Mann Whitney U test was administered to test whether there was a significant difference seen in their FFR parameters. In younger adults, the results revealed that there was a significant difference ($p < 0.001$) in FFR parameters in terms of the fundamental frequency in quiet and noise, the root mean square amplitude in quiet and noise. However in older adults, the top performers and bottom performers in SPIN showed significant difference ($p < 0.05$) in fundamental frequency in quiet and noise and in the V th peak latency. The Table 3.8 shows the results of Mann Whitney U test. Figure 3.15 and 3.16 shows the FFR waveform of an younger adult with top SPIN performer and bottom SPIN performer. Similarly Figure 3.17 and 3.18 shows the FFR waveform of an older adult with top SPIN performer and bottom SPIN performer.

Table 3.8.

The Z value and significance level for latency, amplitude and FFT parameters of brainstem responses between top and bottom SPIN performer in YNH and ONH adults

YNH			**ONH**		
Parameters	**Z value**	**P value**	**Parameters**	**Z value**	**P value**
SPIN	-5.966	.000	SPIN	-5.187	.000
F0 in quiet	-4.263	.000	F0 in quiet	-2.852	.004
F0 in noise	-3.332	.001	F0 in noise	-2.653	.008
rms amplitude in quiet	-2.664	.008	V th latency	-2.878	.004
rms in noise	-3.004	.003			

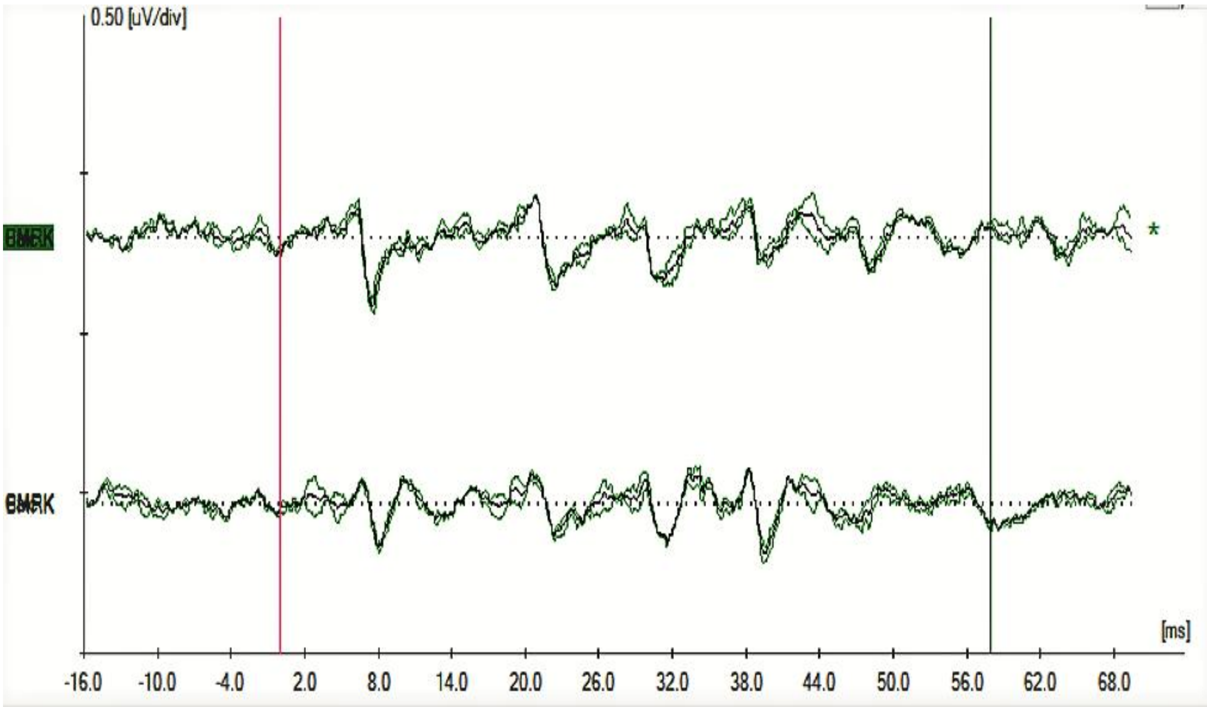


Figure 3.15. FFR of an 28 years female with top SPIN performer wherein the first waveform indicates the FFR in quiet and the bottom waveform indicates the FFR in noise.

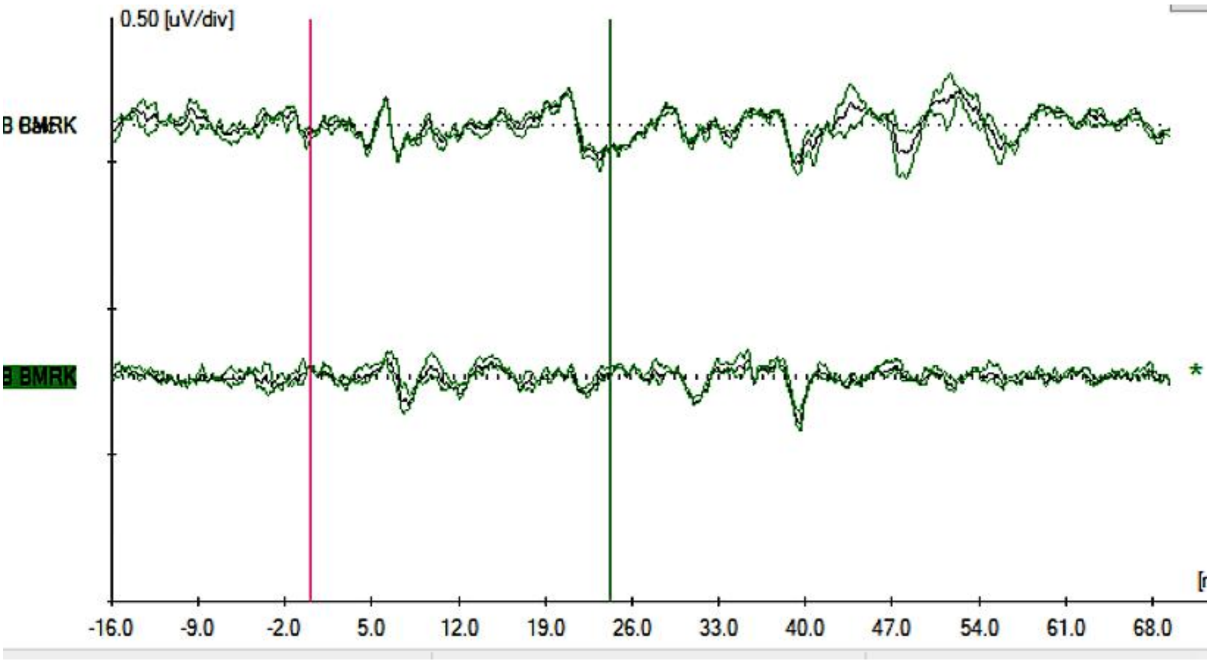


Figure 3.16. FFR of an 32 years female with bottom SPIN performer wherein the first waveform indicates the FFR in quiet and the bottom waveform indicates the FFR in noise.

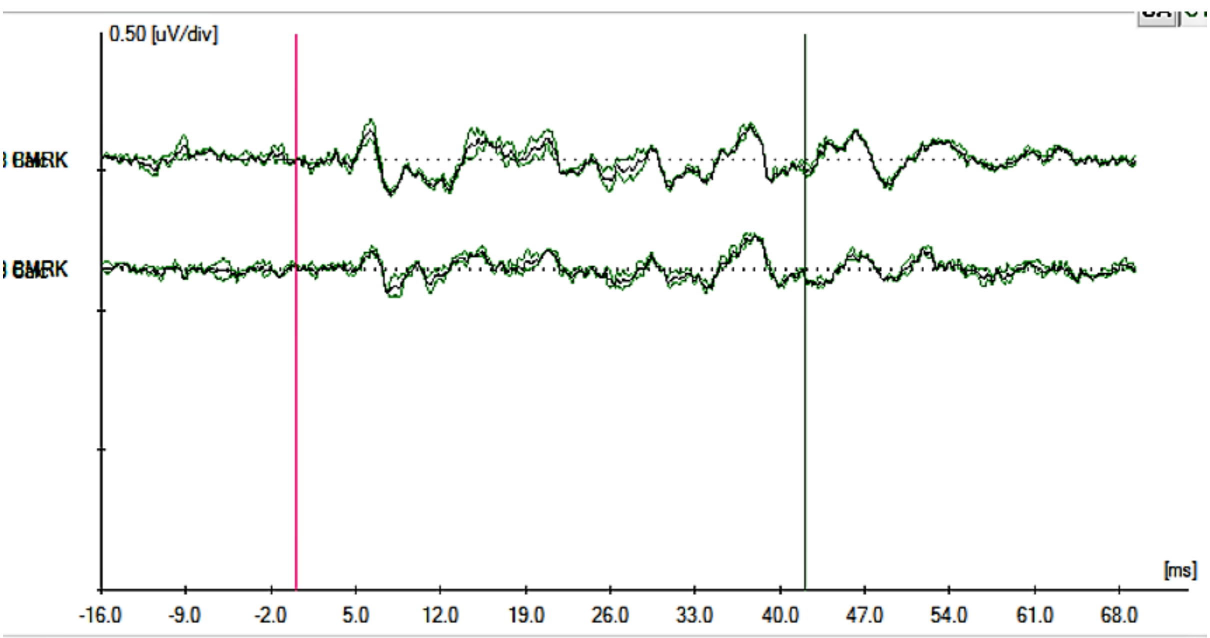


Figure 3.17. FFR of an 63 years female with top SPIN performer wherein the first waveform indicates the FFR in quiet and the bottom waveform indicates the FFR in noise.

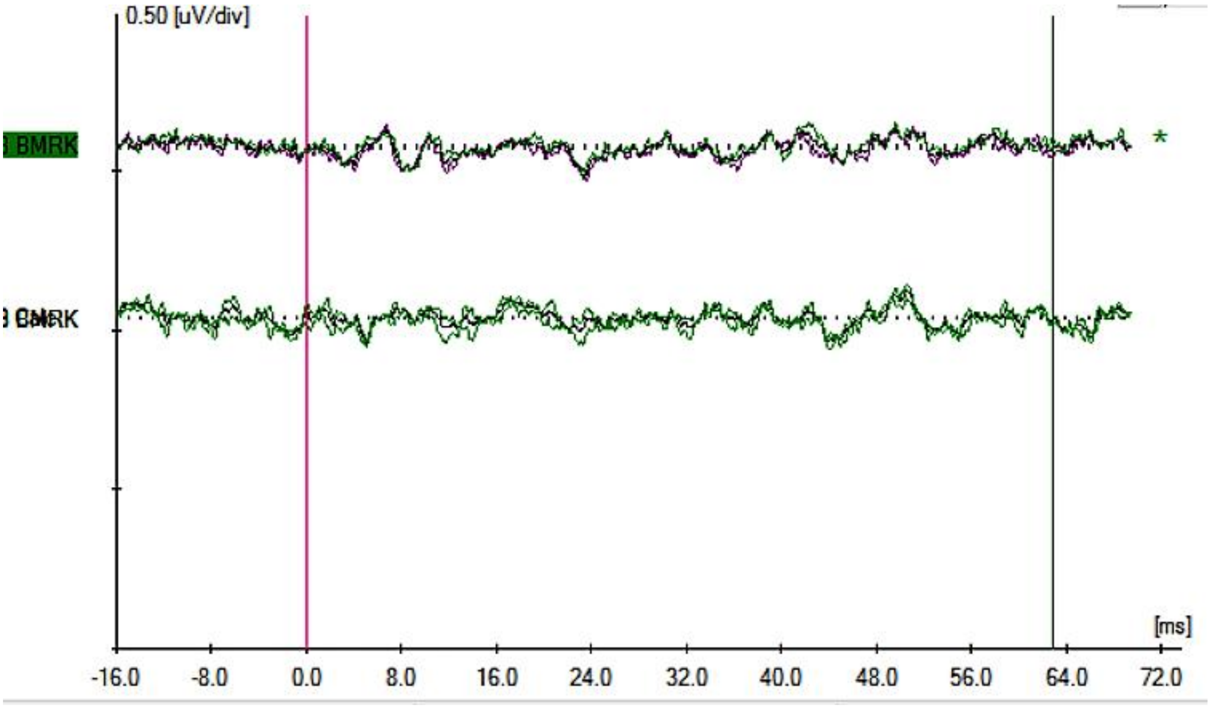


Figure 3.18. FFR of an 62 years male with bottom SPIN performer wherein the first waveform indicates the FFR in quiet and the bottom waveform indicates the FFR in noise.

Overall, the results showed that there is a significant age effect on speech perception in noise and suprathreshold processing. *Table 3.9* shows the summary of Spearmann correlations in YNH and ONH adults with SPIN.

Table 3.9.

Summary of correlation of SPIN with suprathreshold abilities and working memory

Suprathreshold tests	**YNH**	**ONH**
GDT	Strong negative correlation	Strong negative correlation
ITD	Strong negative correlation	Moderate negative correlation
PS and BW	No correlation	Weak correlation
Working memory tests	Forward Span and descending auditory number sequencing- strong positive correlation Backward span and ascending auditory number sequencing- moderate positive correlation	Strong positive correlation for descending and descending auditory number sequencing and no correlation for other working memory tests
Spatial selective attention	No correlation	No correlation
DLF and RND	Moderate negative correlation	Weak correlation
FFR (f0 and f1)	Moderate positive correlation	Moderate positive correlation

Chapter 4

Discussion

One of the major problems faced by older adults is the difficulty to understand speech in the presence of background noise despite of having normal hearing sensitivity. There are many cues that facilitate the perception of speech in background noise like the spectral cues (Fo), the temporal cues and the working memory as per the literature. However, there is lack of studies reported which among these cues are important for the perception of speech in background noise. Aging has a detrimental role which degrades the auditory processes even when the peripheral hearing is normal especially for speech perception in noise (Gordon-Salant & Fitzgibbons, 1993). Hence, the cues that older adults use to understand the speech in noise would be different from a normal hearing young adult. Therefore, the present study incorporated the role of envelope and spectrotemporal cues in understanding speech in noise for old normal hearing adults compared to younger adults. In addition, the role of working memory capacity or the top down processes was also explored.

4.1. Effect of age on cocktail-party listening

The results of the present study showed that there was a significant effect of aging on speech perception in noise As expected, older adults showed poorer performance on speech perception in noise (SPIN) compared to younger adults which suggests that the role of peripheral hearing sensitivity to understand speech in the presence of background noise is limited or in other words, it does not convey information on the perception of speech in the presence of noise (Killon & Niquette, 2000). Further, SPIN is a complex behavior and is influenced by various factors. CHABA (1988) proposed three main reasons to explain poor SPIN in older adults: (1) peripheral, which focused on cochlear damage resulting in impaired audibility and suprathreshold processing skills, (2) central auditory, which included lesion to auditory brainstem and cortical structures and (3) cognitive, which involved age related degeneration in non auditory areas responsible for linguistic and cognitive process.

4.2. Cocktail-party listening and supra-threshold processing

The present results showed that both younger adults and older adults showed strong positive correlations of temporal processing tests such as gap detection (GDT) and Interaural Time Difference (ITD) with SPIN. The results are in agreement with the findings of Helfer and Vargo (2009) who stated that there is a close relationship between the GDT and SPIN. Further, the speech recognition in cocktail party conditions can be improved when there is spatial separation between the speaker and the masker (spatial unmasking) which can be due to head shadow effect improving the signal to noise ratio or due to the difference in time between the target and masker (Zurek, 1993). The GDT and ITD was elevated for older adults which could be due to the reduced neural synchrony which in turn reduces the precision of coding the temporal differences resulting in poor SPIN (Grose & Manro, 2009). Babkoff et al. (2002) also reported that the ITD thresholds were elevated in older individuals for low frequency at low sensation levels. The detriments in the neural synchrony in turn reduces their precision to process the temporal cues especially the binaural cues in turn diminishing their SPIN performance. The strong correlation of the temporal processing tests with SPIN suggests that even though the thresholds for temporal processing are elevated for older adults due to their poor neural synchrony, they still rely on temporal cues for SPIN like the young normal hearing adults to understand speech in adverse listening situations.

However, the results of modulation detection thresholds (detection of changes in the envelope) were in disagreement with these findings of temporal processing. Moore (2008) reported that the complex auditory signal like speech signal with broad frequency spectrum is analysed in terms of its slow varying envelope and temporal fine structure as a series of output from the band pass auditory filters. The modulation detection thresholds assesses the detection of envelops at different modulation frequencies and in the present study the correlation between modulation detection thresholds with speech perception in fluctuating background noise was assessed. Results showed that the peak sensitivity and bandwidth obtained from the modulation detection thresholds at different modulation frequencies had a weak correlation with SPIN. This findings suggests that the envelope cues does not convey information to perceptually segregate the background (especially when the background is fluctuating as in the study when it is multitalker babble) from the target speech though it is sufficient for intelligibility in quiet (Moore, 2008).

The results of the present study also showed that the tests that assess spectral processing (differential limen of frequency and ripple noise discrimination) showed moderate correlation with speech perception in noise in younger adults while there was no correlation in older adults. Each auditory filter shape and the tonotopicity is important for the decoding of the broad band speech stimuli because when the auditory filters are broadened, the coding of the temporal fine structure is altered and varies rapidly (Huss & Moore, 2005) which in turn affects the central mechanism to decode the signal. In older adults, due to the altered filter shape, they do not rely much on this cue for understanding speech in the presence of noise.

4.3. Cocktail-party listening and working memory.

Working memory capacity has an important role in cocktail party listening. The individuals with high working memory capacity are better in inhibiting or ignoring the distracting information which means that it is responsible to maintain attention to the relevant information simultaneously ignoring the irrelevant or the distracting information (Conway &

Engle, 1994). Conway, Cowman and Bunting (2001) demonstrated that the working memory capacity is important for selective attention tasks which are similar to that of cocktail party listening. Working memory is responsible for the manipulation and temporary storage of information which is necessary for complex cognitive tasks. According to Baddley (2000) working memory consists of four components namely central executive, phonological loop, visuospatial sketch pad and episodic buffer. Among these, central executive is an attention control system and the phonological loop is responsible for temporary storage and rehearsing of speech based information which is important for understanding speech in complex situations (Baddley & Hitch, 1974; Baddley, 2000). When there is a mismatch between the speech input and phonological representation in the long term memory, an explicit processing and storage capacity is required to comprehend the input. This mismatch is commonly seen in cocktail party situation, even for individuals with normal hearing. In an unfamiliar situation, individuals tend to depend more on their working memory capacity due to this mismatch indicating that working memory capacity becomes a good predictor for speech perception in noise (Ronnberg et al., 2010; Rudner et al., 2008; Rudner, Foo, Ronnberg, Lunner, 2009).

The results of the present study also showed that both older adults and younger adults with normal hearing utilize their cognitive ability including attention and working memory to understand speech when the speech input is degraded due to presence of background noise. The babble used in the study is predicted to add a perceptual load in turn increasing the need to rely on cognition to understand speech. The age related decline in their working memory capacity could be due to the loss of efficiency in the central and slave systems. This can be attributed to the loss of suprathreshold processing which degrades the speech input (Gregoire & Linden, 1997). However, a strong relationship was observed only between SPIN and descending number sequencing, and the reason being digit span tests do not require greater amount of information processing (Baddley, 1986). In contrast, the descending auditory number sequencing requires more processing and hence require more involvement of central executive. Hence, we can conclude that age related decline can be found in older adults in their working memory capacity. However, not every working memory tests would showed significant correlations with cocktail party effect. A test that requires more informational processing and simultaneous storage which involves more function of the central executive (descending auditory number sequencing) would have strong correlations with cocktail party effect. The present study supports the theoretical model of Ronnberg et al., (2010) which stated that there will be more reliance on working memory capacity when there is a phonological mismatch between the speech input and already stored templates.

4.4. Neural basis for speech perception in noise

The result of the present study showed that there was a significant deterioration in brainstem encoding of speech as indicated by the coding of transient and sustained components with ageing. Further, it was also noted that fundamental frequency served as the major cue to show correlation with speech perception in noise in both the groups irrespective of the age factor and it was the major factor which showed significant difference between top and bottom performers in both younger and older adults. This finding attributes to the general conclusion that F0 serves as a major cue in understanding speech in adverse listening or in multiple speaker condition which help us to identify the speaker and thereby helping to separate the target vs. the masker. Even though the older adults benefit less from the pitch/ F0 cues compared to younger adults (Helfer & Freyman, 2008) which is reflected even in the brainstem encoding it is indicated that F0 remains robust when compared to other parameters and correlates well with SPIN scores (Anderson et al., 2011). The current study revealed that even when the groups matched in their pure tone thresholds and differed only in terms of their age, there was significant deterioration in the SPIN scores for older adults when compared to younger adults. Similarly, both younger and older adults differed in the speech ABR except for the coding of harmonics (F1) and in both the groups showed correlations with SIN scores. Hence, individuals in each group were divided into top and bottom performers based on their SPIN scores and FFR data was analysed. The results showed there was a significant difference between top and bottom performers primarily in terms of their F0 coding in quiet and in noise.

Thus, to conclude YNH adults showed significant correlation between spectral, temporal and selective attention tests with speech perception in noise but older adults showed correlations only between the temporal tests and attention with speech perception in noise. Results also showed that though older adults showed significant difficulty in understanding speech in the presence of multitalker babble or in a cocktail party situation eventhough they had normal hearing sensitivity as that of younger adults. This difference in speech perception in noise between both the groups can be attributed to their poor spectrotemporal coding. YNH adults showed significant correlation with the tests that assess both spectral as well as temporal coding with speech perception in noise. On the other hand older adults, showed significant correlation only with the temporal tests and speech perception in noise which means that despite of having poor temporal resolution they rely on their temporal cues to understand speech in adverse listening conditions which is also being facilitated by their cognitive cues and relies less on their spectral cues which could be because of the poor spectral resolution in them due to widened auditory filters. During testing, many of the older adults found the task of frequency discrimination as very difficult and many of them scored the maximum value (as per the default) in ripple noise discrimination and absolute frequency difference limen (1.83 dB and 101.01 Hz in ripple noise discrimination and DLF respectively). The study supports the hypothesis that the audiogram or hearing thresholds are not good predictors of speech perception in noise especially in multitalker babble. Further, both the groups rely on their cognitive cue to a certain extent to understand speech in adverse listening conditions.

CHAPTER 5

Summary and Conclusion

The present study investigated the effect of age, suprathreshold processing and selective attention on speech perception in noise in younger and older adults. A total of 92 participants with normal hearing sensitivity were included in the study. They were divided into two groups based on their age. 52 young normal hearing adults in the age range of 20 -40 years and 40 older normal hearing adults. Tests administered included speech perception in noise test to assess cocktail party listening, temporal and spectral processing tests to assess suprathreshold processing and auditory digit span, auditory digit sequencing and spatial selective attention to assess working memory. Results showed that older adults performed poorer than younger adults in all the tests. The results also showed that despite of having normal hearing older adults showed poor performance in speech perception in noise which could be due to their age related decline in the suprathreshold processing and top down processes. Among the suprathreshold tests, the temporal tests showed more significant correlation with speech perception in noise rather than spectral ones. This can be attributed to the disrupted neural synchrony which is due to poor frequency selectivity as observed through ripple noise discrimination. Individuals rely more on temporal cues due to poorer frequency resolution and phase locking mechanism and also on top down processes such as working memory. A degraded speech input also lead them to rely more on their higher cognition.

References

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal of Audiology, 47(Suppl 2),* S53–71.

Anderson, S., Parbery-Clark, A., Yi, H. G., & Kraus, N. (2011). A neural basis of speech-in-noise perception in older adults. *Ear and hearing*, *32*(6), 750.

ANSI S3.1-1999. American National Standard Maximum Permissible Ambient Noise Levels for Audiometric Test Rooms.

Babkoff, H., Muchnik, C., Ben-David, N., Furst, M., Even-Zohar, S., & Hildesheimer, M. (2002). Mapping lateralization of click trains in younger and older populations. *Hearing research, 165(1)*, 117-127.

Baddeley, A. D. (1986). *Working Memory*. Oxford: Oxford Scientific Publications.

Baddeley, A. D. (2000). Short-term and working memory. *The Oxford handbook of memory*, 77-92.

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of learning and motivation*, *8*, 47-89.

Basavaraj, V., & Venkatesan, S. (2009). *Ethical Guidelines for Bio-behavioral research involving human subjects*. Mysore: Dr. Vijayalakshmi Basavaraj, Director, All India Institute of Speech and Hearing. ISBN: 978-81-909355-6-2.

Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT, Cambridge.

Cherry, E. C. (1953). Some Experiments on the Recognition of Speech, with One and Two Ears. *Journal of the Acoustical Society of America, 25*, 975-979.

Clinard, C. G., Tremblay, K. L., & Krishnan, A. R. (2010). Aging alters the perception and physiological representation of frequency: evidence from human frequency-following response recordings. *Hearing research*, *264*(1), 48-55.

Conway, A. R., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic bulletin & review*, *8*(2), 331-335.

Conway, A. R., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General*, *123*(4), 354.

Culling, J. F., & Darwin, C. J. (1993). Perceptual separation of simultaneous vowels: within and across formant grouping by F0. *Journal of Acoustical Society of America, 93*, 3454–3467.

Cunningham, J., Nicol, T., Zecker, S. G., Bradlow, A., & Kraus, N. (2001). Neurobiologic responses to speech in noise in children with learning problems: deficits and strategies for improvement. *Clinical Neurophysiology*, *112*(5), 758-767.

Dempster, F. N. (1991). Inhibitory processes: a negleted dimension of intelligence. *Intelligence, 15*, 157–173.

Du, Y., He, Y., Ross, B., Bardouille, T., Wu, X.-H., Li, L., & Alain, C. (2011). Human auditory cortex activity shows additive effects of spectral and spatial cues during speech segregation. *Cerebral Cortex, 21*, 698–707.

Dubno, J. R., Horwitz, A. R., & Ahlstrom, J. B. (2002). Benefit of modulated maskers for speech recognition by younger and older adults with normal hearing. *Journal of the Acoustical Society of America, 111(6)*, 2897–2907.

Finney, D. J. (1952). *Statistical method in biological assay*. London, England: C. Griffen.

Fulton, S., Lister, J., Bush, A., Edwards, J., & Andel, R. (2015). Mechanisms of the hearing–cognition relationship. *Seminars in Hearing, 36(3)*, 140–149.

Galbraith, G. C., Amaya, E. M., de Rivera, J. M. D., Donan, N. M., Duong, M. T., Hsu, J. N., ... & Tsang, L. P. (2004). Brain stem evoked response to forward and reversed speech in humans. *Neuroreport*, *15*(13), 2057-2060.

Galbraith, G. C., Arbagey, P. W., Branski, R., Comerci, N., & Rector, P. M. (1995). Intelligible speech encoded in the human brain stem frequency-following response. *Neuroreport*, *6*(17), 2363-2367.

Gardi, J., Salamy, A., & Mendelson, T. (1979). Scalp-recorded frequency-following responses in neonates. *Audiology*, *18*(6), 494-506.

Gates, G. A., Beiser, A., Rees, T. S., Agostino, R. B., & Wolf, P. A. (2002). Central auditory dysfunction may precede the onset of clinical dementia in people with probably Alzheimer‟s disease. *Journal of the American Geriatric Society, 50*, 482-488.

Gates, G. A., Cooper, J. C., Jr Kannel, W. B., Miller, N. J. (1990). Hearing in the elderly: The Framingham cohort, 1983–1985. Part I. Basic audiometric test results. *Ear and Hearing, 11*, 247–256.

Gordon-Salant, S. (2006). Speech perception and auditory temporal processing performance by older listeners: Implications for real-world communication. *Seminars in Hearing, 27(4)*, 264–268.

Gordon-Salant, S., & Fitzgibbons, P. J. (1993). Temporal factors and speech recognition performance in young and elderly listeners. *Journal of Speech, Language, and Hearing Research, 36*(6), 1276-1285.

Grassi, M., & Soranzo, A. (2009). MLP: a MATLAB toolbox for rapid and reliable auditory threshold estimation. *Behavior research methods, 41*(1), 20-28.

Green, D. M. (1990). “Stimulus selection in adaptive psychophysical procedures”. *Journal of Acoustical Society of America, 87,* 2662–2674.

Green, D. M. (1993). “A maximum-likelihood method for estimating thresholds in a yes-no task”. *Journal of Acoustical Society of America, 93,* 2096–2105.

Greenberg, S., Marsh, J. T., Brown, W. S., & Smith, J. C. (1987). Neural temporal coding of low pitch. I. Human frequency-following responses to complex tones. *Hearing research, 25*(2), 91-114.

Gregoire, J., & Van der Linden, M. (1997). Effect of age on forward and backward digit spans. *Aging, Neuropsychology, and Cognition, 4*(2), 140-149.

Grose, J. H., & Manro, S. K. (2010). Processing of temporal fine structure as a function of age. *Ear and Hearing, 31(6),* 755–760.

Helfer, K. S., Vargo, M. (2009). Speech recognition and temporal processing in middle-aged women. *Journal of American Academy of Audiology, 20,* 264–271.

Helfer, K., & Freyman, R. (2008). Aging and speech-on-speech masking. *Ear and Hearing, 29,* 87–98.

Huss, M., & Moore, B. C. (2005). Dead regions and pitch perception. *The Journal of the Acoustical Society of America, 117*(6), 3841-3852.

Jerger, J., Jerger, S., Oliver, T., & Pirozzolo, F. (1989). Speech understanding in the elderly. *Ear and hearing*, *10*(2), 79-89.

Johnson, K. L., Nicol, T. G., & Kraus, N. (2005). Brain stem response to speech: a biological marker of auditory processing. *Ear and hearing*, *26*(5), 424-434.

Kidd, G., Arbogast, T. L., Mason, C. R., and Gallun, F. J. (2005). “The advantage of knowing where to listen.” *Journal of Acoustical Society of America*, *118*, 3804–3815.

Killion, M., & Niquette, P. (2000). What can the pure-tone audiogram tell us about a patient's SNR loss? *The Hearing Journal*, *53(3)*, 46-53.

Krishnan, A. (2002) Human frequency-following responses: representation of steady-state synthetic vowels. *Hearing Research*, *166*, 192–201.

Krishnan, A., Swaminathan, J., & Gandour, J. T. (2009). Experience-dependent Enhancement of Linguistic Pitch Representation in the Brainstem Is Not Specific to a Speech Context. *Journal of Cognitive Neuroscience*, *21*(6), 1092–1105.

Kumar, U. A., & Sandeep, M. (2013). *Development and Test Trail of Computer Based Auditory-Cognitive Training Module for Individuals with Cochlear Hearing Loss.* Unpublished Departmental Project: AIISH; Mysuru.

Lin, F. R., Ferrucci, L., Metter, E. J., An, Y., Zonderman, A. B., & Resnick, S. M. (2011). Hearing Loss and Cognition in the Baltimore Longitudinal Study of Aging. *Neuropsychology*, *25*(6), 763–770.

Lin, F. R., Yaffe, K., Xia, J., Xue, Q.-L., Harris, T. B., Purchase-Helzner, E., … for the Health ABC Study. (2013). Hearing Loss and Cognitive Decline Among Older Adults. *JAMA Internal Medicine*, *173*(4), 10. 1001/jamainternmed.2013.1868.

Methi, R., Avinash, & Kumar, U. A. (2009). Development of sentencematerial for Quick Speech in Noise test (Quick SIN) in Kannada. *Journal of Indian Speech and Hearing Association, 23(1),* 59–65.

Moore, B. C. (2008). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *Journal of the Association for Research in Otolaryngology*, *9*(4), 399-406.

Moore, B. C., & Moore, G. A. (2003). Discrimination of the fundamental frequency of complex tones with fixed and shifting spectral envelopes by normally hearing and hearing-impaired subjects. *Hearing research*, *182*(1), 153-163.

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology, 11*, 56-60.

Nelson, P. B., Jin, S. H., Carney, A. E., & Nelson, D. A. (2003). Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners. *The Journal of the Acoustical Society of America*, *113*(2), 961-968.

Rönnberg, J., Rudner, M., Lunner, T., & Zekveld, A. A. (2010). When cognition kicks in: Working memory and speech understanding in noise. *Noise and Health*, *12*(49), 263.

Rudner, M., Foo, C., Rönnberg, J. et al. (2009). Cognition and aided speech recognition in noise: Specific role for cognitive factors following nine-week experience with adjusted compression settings in hearing aids. *Scandinavanian Journal of Psychology, 50,* 405–418.

Rudner, M., Foo, C., Ronnberg, J., & Lunner, T. (2007). Phonological mismatch makes aided speech recognition in noise cognitively taxing. *Ear and Hearing, 28(6),* 879–892.

Ruggles, D., & Shinn-Cunningham, B. (2011). Spatial selective auditory attention in the presence of reverberant energy: Individual differences in normal-hearing listeners. *Journal of the Association for Research in Otolaryngology, 12(3)*, 395–405.

Russo, N., Nicol, T., Musacchia, G., & Kraus, N. (2004). Brainstem responses to speech syllables. *Clinical Neurophysiology*, *115*(9), 2021-2030.

Schneider, B. A., Li, L., Daneman, M. (2007). How competing speech interferes with speech comprehension in everyday listening situations? *Journal of American Academy of Audiology, 18,* 559–572.

Skoe, E., & Kraus, N. (2010). Auditory brainstem response to complex sounds: a tutorial. *Ear and hearing*, *31*(3), 302.

Soranzo, A., & Grassi, M. (2014). PSYCHOACOUSTICS: a comprehensive MATLAB toolbox for auditory testing. *Frontiers in psychology*, *5*, 712.

Souza, P., Boike, K., Witherell, K., & Tremblay, K. (2007). Prediction of speech recognition from audibility in older listeners with hearing loss: Effects of age, amplification, and background noise. *Journal of the American Academy of Audiology, 18*, 54–65.

Vander Werff, K. R., & Burns, K. S. (2011). Brain stem responses to speech in younger and older adults. *Ear and hearing*, *32*(2), 168-180.

Zurek, P. M. (1993). "Binaural advantages and directional effects in speech intelligibility". *In Acoustical Factors Affecting Hearing Aid Performance, 2nd ed.*, edited by Studebaker G. A. and Hockberg I. (Allyn and Bacon, Needham Heights, MA), 255–276.