AUDITORY PLASTICITY IN MUSICIANS: A COMPARATIVE STUDY

Register No. 10AUD008

A dissertation submitted in part fulfillment for the degree of

M.Sc., (Audiology)

University of Mysore, Mysore.

ALL INDIA INSTITUTE OF SPEECH AND HEARING,

NAIMISHAM CAMPUS, MANASGANGOTHRI,

MYSORE – 570006.

MAY 2012

CERTIFICATE

This to certify that this dissertation entitled **"AUDITORY PLASTICITY :A COMPARATIVE STUDY"** is the bonafide work in part fulfillment for the degree of Masters in Science (Audiology) of the student (Register No. 10AUD008). This has been carried out under the guidance of a faculty of this institute and has not been submitted earlier to any other university for award of any other diploma or degree.

Mysore
May 2012

Dr. S. R. Savithri
Director
All India Institute of Speech & Herring
Manasgangothri, Mysore
570006

TABLE OF CONTENTS

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

From the cochlea to the auditory cortex, sound is encoded at multiple locations along the ascending auditory pathway, eventually leading to conscious perception (Kraus 2007). Speech is a stream of acoustic elements produced at an astounding average rate of three to six syllables per second (Laver, 1994).The ability to decode these elements in a meaningful manner is a complex task that involves multiple stages of neural processing.

Models examining the neural bases of human speech perception have focused primarily on the cerebral cortex (Bennett & Hacker,2006; Hickok & Poeppel, 2007; Naatanen, 2001; Poeppel &, Poeppel, Idsardi, & van Wassenhove, 2008; Scott & Johnsrude, 2003; Scott & Wise, 2004; Tervaniemi & Hugdahl ,(2003). However, before speech can be perceived and integrated with long-term stored linguistic representations, relevant acoustic cues must be represented through a neural code and delivered to the auditory cortex with temporal and spectral precision by subcortical structures (Eggermont, 2001; Hickok & Poeppel,2007; Poeppel & Hickok, 2004; Poeppel et al., 2008).

Neural plasticity is a term used to describe alterations in the physiological and anatomical properties of neurons in the brain in association with auditory stimulation and deprivation. Depending on the experience, mechanism of plasticity can involve synaptic changes that occur rapidly or slowly over a period of time (Tremblay & Kraus 2001). Everyday learning and training involves of continuous improvement of our abilities the sensory, cognitive & behavioural levels (Menning, Roberts & Pantev 2000). Peripheral and central structures along the auditory pathway contribute to speech processing and learning. However, because speech requires the use of functionally and acoustically complex sounds which necessitates high sensory and cognitive demands, long-term exposure and experience using these sounds is often attributed to the neocortex with little emphasis placed on subcortical structures (Song, Skoe,&Kraus 2008).

Auditory processing is related to language and cognitive function, and impaired auditory processing negatively affects the quality of life of many people. Recent studies suggest that the malleability of the auditory system may be used to study the interaction between sensory and cognitive processes and to enhance human well-being (Kraus & Banai, 2005). Long-term and short-term auditory experiences have been shown to enhance the brainstem responses to complex, behaviourally relevant sounds. Depending on the experience mechanism of plasticity involves synaptic changes that occur rapidly or slowly over a longer period of time (Tremblay & Kraus 2002).

Music is a complex auditory task and musicians spend years fine-tuning their skills. It is no wonder that previous research has documented neuroplasticity to musical sounds as a function of musical experience (Fujioka, Trainor,Ross,Kakigi,Pantev,2005;Koelsch,Schroger, & Tervaniemi, 1999; Musacchia, Sams, Skoe, & Kraus, 2007; Pantev et al., 1998; Pantev, Roberts, Schulz, Engelien, & Ross, 2001; Tervaniemi, Rytkonen,Schroger, Ilmoniemi, & Naatanen, 2001). The domains of music and language share many features, the most direct being that both exploit changes in pitch patterns to convey information. Music uses pitch contours and intervals to communicate melodies and tone centers. Pitch patterns in speech convey prosodic information; listeners use prosodic cues to identify indexical information, i.e., information about the speaker’s intention as well as emotion and other social factors.

Structural brain changes after only 15 months of musical training in early childhood, which were correlated with improvements in musically relevant motor and auditory skills. These findings shed light on brain plasticity, and suggest that structural brain differences in adult experts (whether musicians or experts in other areas) are likely due to training-induced brain plasticity. Listening to music involves both high cognitive demands and auditory acuity; these subcortical enhancements may result from corticofugal (top–down) mechanisms. With long-term musical experience, the musician’s brain has shown functional and structural adaptations for processing. Prior investigations into the neurological effects of musical experience have mainly focused on the neural plasticity of the cortex but recent studies have shown that neural plasticity also extends to the subcortical auditory system.

A theory to account for the interactions between sensory input and top-down processes (e.g., attention, language, and memory) is the Reverse Hierarchy Theory (RHT) (Ahissar and Hochstein, 2004). The RHT postulates that the performance of a perceptual task is first based on the highest available level of sensory representation. If the task cannot be accomplished at that level (because of poor sensory resolution), it proceeds down the representational hierarchy to obtain more detailed, lower-level cues that participate in generating the percept. Because the top-down mechanism was originally proposed for the impact of higher-order visual cortical areas to lower-order cortical areas, the focus has been primarily on the intracortical feedback pathway. Recent studies have extended the RHT to auditory perception (Nahum et al., 2008; Gutschalk et al., 2008), and top-down corticofugal enhancement of brainstem representation of selective features of sound provides evidence for the expansion of the RHT theory outside the cortical areas (Suga, 2008; Luo et al., 2008; Perrot et al., 2006). Recent findings indicate that cortical activation shapes the tuning properties of neurons in the cochlear nucleus (Luo et al., 2008) similar to intracortical, experience-dependent shaping of receptive fields observed in primary auditory cortex (Schreiner and Winer, 2007; Fritz et al., 2007; Atiani et al., 2009).

Kraus & Wong (2007) found more robust and faithful encoding of linguistic pitch information by musicians. Such encoding, arguably associated with increased musical pitch usage, may reflect a positive side effect of context-general corticofugal tuning of the afferent system, implying that long-term music-making may shape basic sensory circuitry.

Musicians have a variety of perceptual and cortical specializations compared to non-musicians. Recent studies have shown that potentials evoked from primarily brainstem structures are enhanced in musicians, compared to non-musicians. Specifically, musicians have more robust representations of pitch periodicity and faster neural timing to sound onset when listening to sounds or both listening to and viewing a speaker. However, it is not known whether musician-related enhancements at the subcortical level are correlated with specializations in the cortex (Musacchia, Strait & Kraus 2008). The effects of musical experience on the nervous system include relationships between brainstem and cortical Evoked Potentials recorded simultaneously in the same subject to seen and heard speech. Moreover, these relationships were related to behavioural measures of auditory perception and were stronger in the audiovisual condition. This implies that musical training promotes plasticity throughout the auditory and multisensory pathways. This includes encoding mechanisms that are relevant for musical sounds as well as for the processing of linguistic cues and multisensory information (Musacchia ,Strait &Kraus 2008).

Hearing speech in noise is a difficult task for everyone, but young children and older adults are particularly vulnerable to the deleterious effects of background noise. Children with learning disorders can exhibit noise exclusion as a primary symptom (Sperling, Lu,Manis, & Seidenberg, 2005). Musicians, in contrast, demonstrate enhanced noise-exclusion abilities (Parbery-Clark,Skoe, & Kraus, 2009; Parbery-Clark, Skoe, Lam,& Kraus 2009). Musical experience enhances the ability to hear speech in challenging listening environments. Speech in Noise performance is a complex task requiring perceptual cue detection, stream segmentation, and working memory. Musicians performed better than nonmusicians in conditions where the target and the background noise were presented from the same source, meaning parsing was more reliant on the acoustic cues present in the stream (Parbery & Kraus 2009).

There is evidence of musical expertise contributing to an enhanced subcortical representation of speech sounds in noise. Musicians had more robust temporal and spectral encoding of the eliciting speech stimulus, thus offsetting the deleterious effects of background noise. Faster neural timing and enhanced harmonic encoding in musicians suggests that musical experience confers an advantage resulting in more precise neural synchrony in the auditory system. These findings provide a biological explanation for musicians’ perceptual enhancement for speech-in-noise (Anderson & Kraus 2010).

NEED FOR THE STUDY

1. The studies have documented better auditory perceptual skills in trained western musicians when compared to non-musicians. So there is a need to execute the same in Indian musicians.

2. There are only very few studies which were done on the relation between brainstem & cortical plasticity in trained musicians, as the experience increases in terms of years of training and practice. So, there is a call for studies in this direction for musicians who have skilled in Carnatic music.

3. To assess effect of music on speech perception in noise.

AIMS OF THE STUDY

- Brainstem correlation of speech in noise perception in musicians.
- To document the auditory plasticity induced by music in musicians on the basis of experience in Carnatic music.
- To compare the brainstem and cortical plasticity in musicians and non musicians.

CHAPTER 2

REVIEW OF LITERATURE

Review of literature reveals that musical training has an effect on both anatomic structures & auditory functions. These functions include –

- Changes in structural & functional aspects of nervous system
- Enhanced brainstem & cortical plasticity.
- Better perception of speech in noise .
- Enhancements in language related skills .
- Enhanced emotional & cognitive process.

Changes in structural & functional aspects of nervous system

Highly trained musicians exhibit anatomical, functional and event-related specializations compared to non-musicians. From an anatomical perspective, musicians have more neural cell bodies (grey matter volume) in auditory, motor and visual cortical areas of the brain (Gaser and Schlaug, 2003) and have more axonal projections that connect the right and left hemispheres (Schlaug et al., 1995). Not surprisingly, professional instrumentalists, compared to amateurs or untrained controls, have more activation in auditory areas such as Heschl’s gyrus (Schneider et al., 2002) and the planum temporale (Ohishi et al., 2001) to sound. Musical training also promotes plasticity in somatosensory regions; with string players demonstrating larger areas of finger representation than untrained controls (Elbert et al., 1995).

Several studies show differences between the brain of adult musicians and non-musicians. For example, structural MRI studies indicate differences in gray matter between musicians and non-musicians in motor, auditory, and visual brain regions (Gaser et al., 2003). Heschl's gyrus, containing primary auditory area, was found to be larger in musicians than non-musicians and its size correlated with musical proficiency (Schneider et al., 2002). Furthermore, the left planum temporale, which is important for the processing of complex sounds, is relatively larger than the right planum temporale in professional musicians, especially those with absolute pitch (Schlaug 2001). With respect to the integrity of organized neural fibers, white matter tracts also appear to differ between pianists and non-musicians, particularly in a pathway from primary motor cortex to the spinal cord and in a region near Broca's area, which is important for complex aspects of language and music processing (Bengtsson et al., 2005; Han et al., 2009).

At a functional level, the brain responses of adult musicians and non-musicians also differ as measured by Electroencephalography and Magnetic Encephalography. For example, some event related potential responses from auditory cortical areas are larger in musicians compared to non-musicians such as N1 occurring at about 100 ms after stimulus onset, N1c, occurring at about 140 ms and larger in the right hemisphere, and P2, occurring at about 170 ms after stimulus onset (Pantev et al., 1998; Shahin et al., 2003).

Recent studies have suggested that playing a musical instrument "tunes" neural activity peripheral to cortical structures (Musacchia et al., 2007; Wong et al., 2007). These studies showed that evoked responses thought to arise predominantly from brainstem structures were more robust in musicians than in non-musician controls. The observed musician-related enhancements corresponded to stimulus features that may be particularly important for processing music. One such example is observed with the Frequency Following Response (FFR), which is thought to be generated primarily in the inferior colliculus and consists of phase-locked inter-spike intervals occurring at the fundamental frequency (F0) of a sound (Hoormann et al., 1992; Krishnan et al., 2005). Because F0 is understood to underlie the percept of pitch, this response is hypothesized to be related to the ability to accurately encode acoustic cues for pitch. Enhanced encoding of this aspect of the stimulus would clearly be beneficial to pitch perception of music. Accordingly, the previous studies demonstrated larger peak amplitudes at F0 and better pitch tracking in musicians relative to non-musicians. Another example was observed with wave delta (~8ms post-acoustic onset) of the brain stem response to sound onset, which has been hypothesized to be important for encoding stimulus onset (Musacchia et al., 2006,2007). Stimulus onset is an attribute of music important for denoting instrument attack and rhythm, and therefore it is perhaps not surprising that the authors observed earliest wave delta responses in musicians than non-musicians. More importantly, FFR and wave delta enhancement in musicians was observed with both music and speech stimuli and was largest when subjects engaged multiple senses by simultaneously lip reading or watching a musician play. This suggests that while these enhancements may be motivated by music related tasks, they are pervasive and apply to other stimuli which possess those stimulus characteristics.

Musacchia, Strait & Kraus (2008) studied the relationship between evoked potentials and musical experience. They recorded simultaneous brainstem and cortical evoked potentials (EP) in musicians and non-musician controls. Because previous research has shown that musician related effects extend to speech and multi-sensory stimuli, the speech syllable /da/ was presented in three conditions: when subjects listened to auditory sound alone, when the subjects simultaneously watched a video of a male speaker saying /da/ and when they viewed the video alone. The analysis focused on comparing measures of the speech evoked brainstem response that have been previously reported as enhanced in musicians with well established measurements of cortical activity (e.g., P1-N1-P2 complex). The first picture that emerged from the data is that recent musical training improves one's auditory memory and shapes composite P1-N1 and pitch encoding (F0). The EP and behavior correlations suggest that complex auditory task performance is related to the strength of the P1-N1 response. The instrumental musicians performed better in the behavioral tests and had steeper P1-N1 slopes than non-musicians. With regard to evoked potentials thought to arise primarily from cortical structures, musicians show enhancements of the P1-N1-P2 complex to pitch, timing, and timbre features of music, relative to non-musicians (Pantev et al., 2001).

However, it was not only the individual tests and measures that were music related. Musicians had a statistically stronger correlation between this set of brain and behavior measures than non-musicians. While it is well known that trained musicians outperform untrained controls and have more robust evoked-potentials than non-musicians, the previous data showed that the accord, or relationship, between brain and behavior is also improved in musicians.

In recent years, musicians have been used as a model for experience induced plasticity, which is known to be expressed in Auditory Evoked Potentials (AEP) in adults (Trembly et al., in 2001). Shahin, Roberts & Trainor (2004) compared AEPs evoked by pure tones, violin and piano tones in young 4- to 5- year old children with age matched non-musician children. The aim of the study was to assess whether AEP components are sensitive to musical experience at this age and, if so, which components are affected. Before conducting the main study AEP responses in independent cohorts of non-musician children between 4 and 15 years of age to the same tones was observed. Larger amplitude P1, N1, and P2 responses were found in 4-to 5-year-old musically experienced children compared with musically less experienced children. Furthermore, the P2 enhancement was specific to the instrument of practice. Thus AEPs differ between musical and control children as young as 4 years of age, and the differences reflect specific musical experience. Comparison of piano-evoked N1 and P2 responses in 4- to 5-year-old musicians (most of whom were pianoists) was done. The cross sectional findings suggest that musical experience may have advanced the developmental trajectory for sounds of the instrument of training.

For a sequential stimuli, occasional wrong notes in a short melody that is repeated in different keys (i.e., starting on different notes) from trial to trial, elicit frontally negative event-related potential called mismatch negativity (MMN). While MMN to such melodic changes was present in both musicians and non-musicians, it was much larger in musicians (Fujioka et al., 2004). In terms of polyphonic music, changed notes in either of the simultaneous melodies elicit MMN responses that are larger in musicians than non-musicians (Fijioka et al., 2005).

The research done on musicians has revealed the advantages in different aspects when compared to non-musicians. Studies have reported that music training can not only improve the skills related to music perception, but also other different aspects like improvement in linguistic skills, working memory, temporal abilities, perception of emotions and also ability to perceive speech in the presence of noise.

Enhanced brainstem & cortical plasticity.

Ample literature exists to address neural encoding of speech sounds from the eighth nerve (Delgutte, 1980; Sachs and Young, 1980; Miller and Sachs, 1983, 1984), cochlear nucleus (Caspary et al., 1977; Palmer et al., 1986; Keilson et al., 1997; Rhode, 1998; Recio and Rhode, 2000), and brainstem (Galbraith et al., 1995, 1997; Krishnan, 1999, 2002). Evoked potentials were used to analyze the development of the auditory brainstem response to click and speech sounds in children between the ages of 3 and 12 years. The neural response to a click stimulus showed similar response timing across all age groups, in agreement with previously established reports (Salamy, 1984; Gorga et al., 1989; Ponton et al., 1992; Abdala and Folsom, 1995; Hurley et al., 2005). In contrast, peak latency measurements throughout the brainstem response to speech were significantly later for 3- to 4-year-old children compared with 5-12 years old. Systematic age related changes in the latency of speech evoked binaural interaction component (BIC) were noted. Latency of BIC of speech obtained in children in age range between 6.11 yrs & 7.11 yrs were significantly prolonged compared to children in the age range between 8 to 12 years, whereas there was no difference in latency of BIC for clicks. Prolonged latency of BIC for speech stimulus in age range 6.11 yrs & 7.11 yrs indicates that BIC continues to devlop till 8 years of age. (Sonitha, 2011). Experience-dependent plasticity in humans is derived from literature on statistical learning. The contribution of statistical learning in data, the literature describes a manner with which the auditory system reacts to frequently occurring sounds. At the level of IC, neural populations rapidly adjust their firing patterns based on the statistical distribution of the sounds encountered, and these adjustments improved coding accuracy for sounds occurring most commonly (Dean et al., 2005), even in an on-line manner. (Johnson , Nicol, Zecker, Kraus 2008). Krishnan et al. (2005) found that native Mandarin speakers had increased accuracy in pitch tracking compared to native English-speaking adults, and Musacchia et al. (2007) and Wong et al. (2007) found enhanced brainstem encoding of the F0 in musicians. These studies can only speak to the effect of long-term auditory experiences initiated in childhood. Moreover, short-term training has been shown to improve brainstem timing in children with learning problems (Russo et al., 2005).

Highly skilled violinists and pianists and nonmusician controls listened under conditions of passive attention to violin tones, piano tones, and pure tones matched in fundamental frequency to the musical tones. Compared with nonmusician controls, both musician groups evidenced larger N1c (latency, 138msec) and P2 (latency, 185msec) responses to the three types of tonal stimuli. As in training studies with nonmusicians, N1c enhancement was expressed preferentially in the right hemisphere, where auditory neurons may be specialized for processing of spectral pitch (Shahin, Bosnyak, Trainor, Roberts 2003).

Better perception of speech in noise

Musicians, as a consequence of training that requires consistent practice, online manipulation, and monitoring of their instrument, are experts in extracting relevant signals from the complex soundscape (e.g., the sound of their own instrument in an orchestra). Literature shows that the effect of musical experience is transferred on the skills that subserve successful perception of speech in noise & beyond. A recent Kraus lab study found a distinct speech-in-noise advantage for musicians, as measured by standardized tests of hearing in noise (HINT, Hearing in- noise test; QuickSIN) (Parbery-Clark, Skoe, Lam et al., 2009). Across all participants, the number of years of consistent practice with a musical instrument correlated strongly with performance on QuickSIN, auditory working memory and frequency discrimination. These correlations strongly suggest that such practice fine tunes cognitive and sensory abilities, leading to an overall advantage in speech perception in noise in musicians. The results from the study suggest that musical experience enhances the ability to hear speech in challenging listening environments. SIN performance is a complex task requiring perceptual cue detection, stream segregation, and working memory. Musicians performed better than non-musicians in conditions where the target and the background noise were presented from the same source, meaning parsing was more reliant on the acoustic cues present in the stream.

SIN perception may also be affected by changes in central auditory processing. Aging affects the ability to process pitch cues (Helfer & Vargo 2009). Ability to perceive speech in the presence of the noise in all the three SNRs (0 dB, -5 dB & -10 dB) is better as the experience of the musicians increased. It was found that as the experience of musician increased the ability to perceive speech in the presence of background noise also increased, especially at lower SNRs (Thomas A.O. 2011).

In order to find the effect of musical experience on the neural representation of speech in noise, Parbery-Clark, Skoe & Kraus (2009) compared sub-cortical neurophysiological responses to speech in quiet and noise in a group of highly trained musicians and non-musician controls. Speech evoked auditory brainstem responses for speech syllable /da/ indicated that musicians exhibited more responses in background noise than control group. Also, musicians had earlier response onset timing, as well as greater phase locking to the temporal waveform and stimulus harmonics, than non-musicians. They also found that earlier response timing and more robust brainstem responses to speech in background noise were both correlated to better speech in noise perception as measured through HINT. They concluded that musical experience resulted in more robust subcortical representation of speech in the presence of background noise, which may contribute to musician’s behavioural advantage for speech in noise perception. Musicians also exhibited more robust responses to the steady state portion of the stimulus in the presence of background noise. By calculating the degree of similarity between stimulus waveform and the sub cortical representation of the speech sound, it was found that musicians had higher stimulus-to-response correlations in noise than non-musicians. Greater stimulus to response correlation is indicative of more precise neural transcription of stimulus features. One possible explanation for this musician enhancement in noise may be based on Hebbian principle, which posits that the associations between neurons that are simultaneously active are strengthened and those that are not are subsequently weakened (Hebb, 1949). It is speculated that extensive musical training may lead to greater neural coherence. This strengthening of the underlying neural circuitry would lead to a better bottom-up, feed forward representation of the signal.

It is well documented that the auditory cortex sharpens the subcortical sensory representations of sounds through the enhancement of the target signal and the suppression of irrelevant competing background noise via the efferent system (Suga et al., 1997; Zhang et al., 1997; Luo et al., 2008). The musician's use of fine grained acoustic information and lifelong experience with parsing simultaneously occurring melodic lines may refine the neural code in a top-down manner such that relevant acoustic features are enhanced early in the sensory system. This enhanced encoding improves the subcortical signal quality, resulting in more robust representation of the target acoustic signal in noise. The subcortical encoding of the $F0$ is an important factor in SIN perception. The $F0$ and other pitch cues contribute to auditory object identification, allowing the listener to "tag" the target voice with a specific identity and to follow this particular voice from among competing voices or other noises. The ability to distinguish between competing streams of information is dependent, in part, on the $F0$, as demonstrated by enhanced discrimination of vowels with greater $F0$ separation between concurrent vowels (Assmann & Summerfield 1987; Culling & Darwin 1993) and sentences (Brokx & Nooteboom 1982; Bird & Darwin 1998).

The improved stimulus to response correlation in the noise condition was related to greater neural representation of the higher harmonics but not the fundamental frequency in noise. Musicians, through the course of their training, spend hours producing, manipulating, and attending to musical sounds that are spectrally rich. The spectral complexity of music is partially attributable to the presence and relative strength of harmonics as well as the change in harmonics over time. Musicians have enhanced cortical responses to their primary instrument suggesting that their listening and training experience modulates the neural responses to specific timbres (Pantev et al., 2001; Margulis et al., 2009).

Enhancements in language related skills

The domains of music and language share many features, the most direct being that both exploit changes in pitch patterns to convey information. Music uses pitch contours and intervals to communicate melodies and tone centres. Pitch patterns in speech convey prosodic information; listeners use prosodic cues to identify indexical information, i.e., information about the speaker's intention as well as emotion and other social factors. Further, in tonal languages, changes in pitch are used lexically, ie, in differentiating between words (e.g., Mandarin Chinese: ma high level 'mother', ma high rising 'hemp' , ma low falling rising, 'horse', ma high falling 'scold') .

A significant body of research has focused on the extent to which musical experience provides benefits in language abilities; the results unambiguously suggest that musicians show enhanced processing of prosodic and linguistic pitch. Musicians show an enhanced ability to detect subtle incongruity in prosodic pitch as well as consistent neural differences relative to nonmusicians (Besson, Schon, Moreno, Santos & Magne, 2007; Magne, Schon, & Besson, 2006). Differences between musicians and nonmusicians show up even during pre-attentive stages of auditory processing (Krishnan et al., 2009; Musacchia et al., 2007; Wong & Perrachione, 2007). Frequency following responses (FFR), which ensemble neural responses originating at the auditory brainstem that reflect phase-locking to stimulus features, were recorded from musicians and non-musicians who were listening to the speech syllable /da/ (Musacchia et al., 2007). Relative to non-musicians, musicians showed more robust encoding of timing and pitch features in the speech signal at the level of the brainstem. Using FFR as an index, musicians showed a superior representation of dynamic pitch contours, as reflected by improved pitch tracking accuracy at the level of brainstem (Wong et al., 2007). Experience with one’s native language shapes not only speech perception but auditory processing in general. Thus, native speakers of Mandarin (in which pitch provides meaningful information) were better at processing pitch contours even in a nonlinguistic context, compared to native speakers of English (Bent, Bradlow, & Wright, 2006). At the physiological level, Mandarin speakers show more robust encoding of the pitch content of Mandarin sounds at cortical and subcortical levels of their auditory system, suggesting that language experience fundamentally changes the neural circuitry of the auditory pathway (Krishnan, Xu, Gandour, & Cariani, 2005). The ability to track non-native pitch contours correlated positively with number of years of musical training, suggesting that it was musical experiences that improved lower level representation of non-native pitch. Using synthetic speech stimuli that contain F0 contours representative of citation forms of Mandarin and Thai lexical tones, the major finding of this study demonstrates that experience-dependent brainstem mechanisms for pitch representation, as reflected in pitch-tracking accuracy and pitch strength, are more sensitive in tone (Chinese, Thai) than non-tone (English) language speakers. (Krishnanan 2009). Findings of Chandrasekaran 2009 suggest that musicians showed superior cortical representation of linguistic pitch in a non-native language relative to non-musicians. In their study, native tone-language speakers showed the strongest representation of pitch, suggesting that the context of long term training matters. From a functional perspective, the enhanced cortical and brainstem representations are indeed relevant. Musicians showed a superior propensity to use pitch in lexical contexts during a language learning task, relative to non-musicians (Wong & Perrachione, 2007). Musician’s enhancement is not just restricted to pitch features. Studies also have demonstrated that musicians show superior brainstem representation of timing and harmonic structure in speech, features that are important for differentiating speech sounds (Musacchia et al., 2007;

Parbery-Clark, Skoe, et al., 2009). Taken together these studies demonstrate that musicians show a distinct advantage in the early auditory processing of speech features.

In a hallmark study, Chan and colleagues showed that participants with music training exhibited superior verbal memory relative to non-musicians, as indicated by greater number of words recalled in a list learning task (Chan et al., 1998). Children who received instrumental training not only showed enhanced processing of skills related to music, but also showed enhanced vocabulary relative to untrained controls (Forgeard, Winner et al., 2008). In typically developing children with normal reading ability, musical discrimination skills significantly predicted phonological and reading skills (Forgeard, Schlaug et al., 2008).

Effect of Music Training on Emotional and Cognitive Processing

Perception of emotion in speech and music relies on shared acoustic and neural mechanisms (Nair et al., 2002), suggesting that extensive experience in one domain may lend perceptual benefits to the other.

Examining the subcortical encoding of a complex for emotionally salient stimulus (a Child's cry) as a function of music experience, a recent study demonstrated increased neural efficiency in musicians (Strait et al., 2009; Strait, Kraus, Skoe & Ashley, 2009). In this study they aimed to provide a biological basis for musician's enhanced perception of emotion in speech by investigating the contribution of subcortical mechanisms to the processing of vocally communicated emotional states. 30 musicians were included in the study, who were classified into 2 groups based on 2 criteria: musicians by onset age (MusAge) and musicians by years (Mus Yrs). MusAge subjects had begun musical training at or before age of 7 years, whereas Mus Yrs subjects had received more than 10 years of consistent musical experience. Integrity of auditory brainstem was assessed using auditory brainstem responses with both click and speech (/da/). The authors suggested that musical experience has more pervasive domain-general effects on the auditory system than previously documented, resulting in fine neural timing to acoustic features important for vocal communication. The results thus provide evidence for initial biologic involvement of subcortical mechanisms in the auditory processing of communicated states of emotion.

Relative to non-musicians, musicians showed superior encoding of the most acoustically complex portion of the emotional stimuli, consistent with behavioral studies demonstrating enhanced emotional perception in musicians (Thompson, Schellenberg & Husain, 2004). Number of studies have evidenced a musician enhancement for auditory working and verbal memory. While some research has reported musician enhancements for only auditory and not visual working memory, others have found enhancements for both auditory and visual memory. It appears that musical training may have distinct effects on working memory abilities at different stages of development, with musically trained children demonstrating superior verbal and non-verbal working memory but musically trained adults demonstrating only superior verbal working memory. Music training also has been shown to improve working memory (Forgeard, Winner et al., 2008; Jakobson, Lewycky, Kilgour, & Stoesz, 2008; Parbery-Clark, Skoe, Lam et al., 2009; and executive function abilities (Bialystok & DePape, 2009). Musicians are also significantly better than non-musicians in auditory stream segregation, presumably due to their music training (Beauvois & Meddis, 1997; Zendel & Alain, 2009). Music training also has been shown to improve working memory (Forgeard, Winner, et al., 2008; Jakobson, Lewycky, Kilgour, & Stoesz, 2008; Parbery-Clark, Skoe, Lam, et al., 2009), attention (Strait et al., 2010;Tervaniemi et al., 2009), and executive function (Bialystok & DePape, 2009) abilities.

The research done on musicians has revealed the advantages in different aspects when compared to non-musicians. Studies have reported that music training can not only improve the skills related to music perception, but also other different aspects like improvement in linguistic skills, working memory, temporal abilities, perception of emotions and also ability to perceive speech in the presence of noise.

CHAPTER 3

METHOD

The present study aimed to find out the effect of musical training on auditory plasticity and speech perception in noise in musicians with various years of Carnatic vocal musical training or practice, using Speech evoked Auditory Brainstem Response, Speech evoked Late Latency Response, and Speech Perception in Noise (SPIN) tests.

Participants

A total of 50 subjects aged between 7-18 years. 25 children enrolled for Carnatic music learning & 25 untrained children were included in this study. The musicians were classified in to 3 groups.

25 trained musicians were classified as follows :

Group 1: Music learning age ranging from 7-10 yrs with minimum experience of 2- 3 years (5 subjects).

Group 2: Music learning age ranging from 10-13 yrs with minimum experience of 4-5 years (10 subjects).

Group 3: Music learning age ranging from 13-18 yrs with experience of greater than 6 years (10 subjects).

Inclusion Criteria

All the subjects who participated in the present study met the following criteria:

- Normal air conduction and bone conduction thresholds (≤15 dB HL) at all octave frequencies from 250 Hz to 8000 Hz.
- Normal middle ear function (‘A’ type tympanogram at 226Hz probe tone with normal acoustic reflexes in both ears.)
- Speech Recognition Threshold of ±12 dB (re. PTA of 0.5, 1 and 2 KHz)
- Speech Identification Scores of > 90% at 40 dB SL (re. SRT) in both ears.
- No indication of Retrocochlear Pathology(RCP)
- No history of neurological or Otological problems.
- No illness on the day of testing.
- All were native Kannada speakers.
- All were professionally trained in Carnatic vocal music for a duration of minimum 2-3 year.

Environment

All testing was carried out in a sound treated double room situation as per the standards of ANSI S3.1 (1991).

Instrumentation

The following instruments were used in the present study:

1. Orbiter 922 (Madsen Electronics, Denmark), two channel audiometer, calibrated as per ISO 389, with supra-aural headphones (Telephonics TDH39) housed with MX-41/ AR ear cushions with audio cups and a bone vibrator (Radioear B71) were used to assess the pure tone threshold, and for Speech Perception in Noise.
2. GSI Tympstar (Grason- Stadler Inc, USA) middle ear analyzer was used for tympanometry and reflexometry.
3. A laptop (was used to deliver the stimulus for SPIN, which were routed through audiometer.

Stimuli

Recorded phonemically balanced (PB) word list in Kannada developed by Yathiraj and Vijayalakshmi (2005), was used for Speech Perception in Noise (SPIN) Test. It consists of 100 words divided into 4 lists (each containing 25 words).

Procedure

Pure tone Audiometry

Air conduction thresholds for octave frequencies from 250 Hz to 8000 Hz and bone conduction thresholds for octave frequencies from 250 Hz to 4000 Hz were obtained with modified version of Hughson Westlake procedure (Carhart & Jerger, 1959).

Speech Audiometry:

Kannada Spondee words (Rajashekar, B, 1976) were used to obtain the Speech Recognition Threshold (SRT) from both ears. A set of 3 spondees were presented at 20 dBSL with reference to PTA and the minimum level at which the subject correctly identified 2 out of 3 spondees were considered as SRT.

Speech Identification Scores in quiet for both ears were obtained with Kannada PB words (Yathiraj & Vijayalakshmi, 2005). PB words, recorded in the voice of a typical Kannada female speaker were presented to both ears separately at 40 dB SL with reference to SRT. A total of 25 words were presented to each ear separately. Each word was given a score of 4 % and the speech identification scores for each ear separately were calculated in percentage.

Immitance Audiometry:

Immitance Audiometry was carried out with GSI Tympstar (Grason- Stadler Inc, USA) middle ear analyzer using 226 Hz probe frequency. Ipsilateral and contra lateral reflexes were measured for 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz.

Speech Perception in Noise (SPIN)

Speech Perception in Noise test was done using the phonemically balanced (PB) Kannada word list (Yathiraj and Vijayalakshmi, 2005), recorded in the voice of a typical female Kannada speaker. The stimuli were played in a laptop and were routed through the audiometer. The presentation level was 40 dB SL (with reference to SRT) or at most comfortable level. The monosyllables and the speech noise were presented monaurally at two different SNRs (0dB,and -5 dB). 25 monosyllables were presented for each trial. The subjects' task was to perceive the monosyllables presented in the presence of noise and repeat them back. Each word was given a score of 4 %. Number of correctly identified word at different SNRs was noted down to find the SPIN score.

Speech evoked Auditory brainstem response

Biologic Navigator Pro EP System version 7.0 was used for recording speech evoked auditory brainstem response.

Test environment

All the tests were carried out in well illuminated air conditioned rooms with noise levels within permissible limits (ANSI-S.3; 1991).

Test stimulus

The /da/ stimulus is a 40 ms synthesized speech syllable produced using KLATT synthesizer (Klatt, 1980) which is available in the Biologic Navigator Pro EP system in the BIOMARK protocol. This stimulus simultaneously contains broad spectral and fast temporal information characteristic of stop consonants, and spectrally rich formant transitions between the consonant and the steady-state vowel. The fundamental frequency (F0) of the /da/ stimulus linearly rises from 103 to 125 Hz with voicing beginning at 5 ms and an onset noise burst during the first 10 msec. The first formant (F1) rises from 220 to 720 Hz, while the second formant (F2) decreases from 1700 to 1240 Hz over the duration of the stimulus. The third formant (F3) falls slightly from 2580 to 2500 Hz, while the fourth (F4) and fifth formants (F5) remain constant at 3600 and 4500 Hz, respectively. Figure -1.1 shows the time domain waveform of the stimulus and Figure – 1.2 shows the spectral waveform of /da/ stimulus used in the present study.

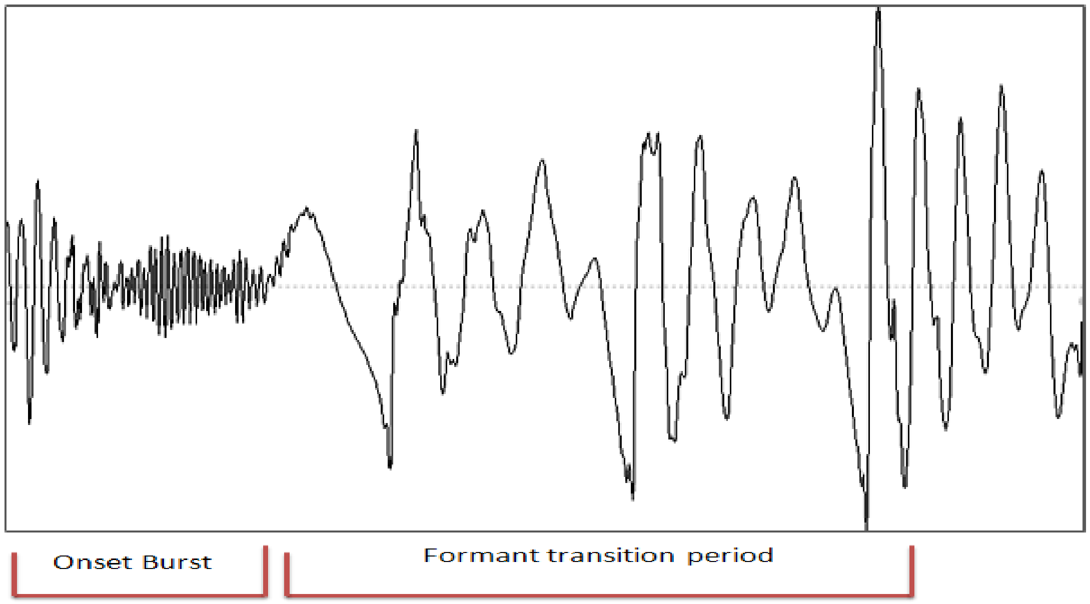

Figure –3.1 Time domain waveform of /da/ stimulus

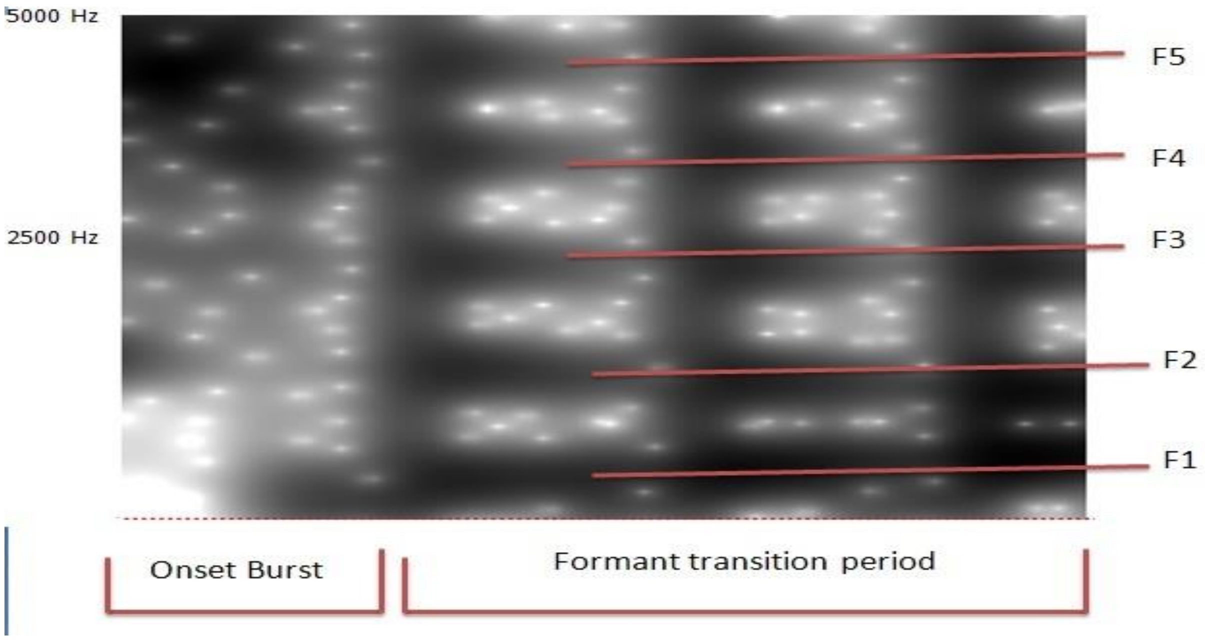

Fig. 3.2 Spectral waveform of /da/ stimulus

Test procedure

During the ABR testing (both clicks and speech- evoked), the subjects were instructed to sit comfortably maintaining a relaxed posture on a reclining chair facing away from the instrument. They were instructed to avoid movement of head, eyes, neck and limbs during testing to avoid artifacts. A muted cartoon video was played in front of the child to reduce the extraneous movements and activity levels.

Electrode placement

Initially the electrode sites were cleaned using skin preparation gel (nuPrep). The gold plated disc type electrodes were placed on the scalp at electrode placement site with adequate amount of ten- 20 conduction paste. The electrodes were secured in place using surgical plaster. The testing was done monaurally. The parameters used to record ABR, which is same as that used by Krizman et al., (2010), is shown in Table 3.1 .

Table 3.1 *Protocol for recording auditory brainstem responses*

Parameters	Target setting for Speech evoked ABR
Stimulus	/da/
Duration	40 ms
Polarity	Condensation & Rarefaction
Stimulus Intensity	80dBSPL
Repetition Rate	10.9
Mode	Ipsilateral
Analysis Time	64 msec including prestimulus period of 11msec
Band Pass Filter	100 to 3000Hz
Electrode Montage	Electrode Montage- Inverting- M1 (Test ear mastoid)

	Non Inverting- Fz (Fore- head) Ground- M2 (Non- test ear mastoid)
Sweeps	3000
Transducer	Biologic Insert
Inter-Electrode Impedance	<2 Kilo Ohms
No. of Channels	One
No. of Replications	Two

Waveforms were collected for rarefaction & condensation polarities and weighted addition was done to obtain calculated waveforms.

The speech evoked ABR and FFR waveform, were converted into ASCII format using the software called ‘AEP TO ASCII’.

Speech evoked LLR

Speech evoked LLR carried out in a sound treated room where the noise levels were as per the guidelines in ANSI S 3.1 (1991). The clients were seated in a reclining chair. The skin surface at the two mastoids (M1, M2), and forehead (Fz) was be cleaned with skin abrasive, to obtain skin impedance of less than 5K ohms for all electrodes. The electrodes were placed with the help of skin conduction paste and surgical plaster was used to secure them tightly in the respective places.

The stimulus and acquisition parameters used for recording LLR are given in Table 3.2.

TABLE 3.2 *Protocol for recording late latency responses*

Parameters	Target setting for LLR
Stimulus	/da/
Duration	40 ms
Polarity	Alternating
Stimulus Intensity	80dBSPL
Repetition Rate	1.1/sec
Analysis Time	500
Band Pass Filter	1 to 30Hz
Electrode Montage	Inverting- M1 Non Inverting- Fz Ground- M2
Sweeps	300

Transducer	Biologic Insert
Impedance	<2 Kilo Ohms

DATA ANALYSIS

Speech evoked ABR is composed of the transient and the sustained responses (also known as frequency following responses). Transient response consists of peak V and peak A whereas the sustained responses consist of peaks D, E, F, and O.

In the present study latency of both the transient as well as sustained responses were analyzed.

1. The transient response was analyzed in terms of latency and amplitude of V and A peak for three repetition rates.
2. The FFR response was analyzed in terms of latency and amplitude of D, E, F, O peaks for the earlier mentioned three repetition rates (the distance between the peak D, E, F, and O is approximately 10msec which gives the information regarding the encoding of fundamental frequency).
3. The sustained portion was analyzed using Fast Fourier Transformation (FFT) for the latency range of 11.4 msec to 40.6 msec for speech evoked ABR to extract the information regarding the coding of fundamental frequency, first formant frequency and second formant frequency at different repetition rates using the MATLAB software.
4. LLR was assessed for P1, N1, and N2 in terms of latency.

Procedure for FFT analysis

To know the coding of fundamental frequency, first formant frequency and higher harmonics, a FFT analysis of the sustained response of the speech evoked ABR was done. This was executed using the MATLAB version 7.0 software (Brainstem toolbox) developed by Kraus (2004) at Northwestern university. For measuring the fundamental frequency and higher harmonics, Fourier analysis was performed on the 11.4–40.6 ms epoch of the FFR in order to assess the amount of activity occurring over three frequency ranges. Activity occurring in the frequency range of the response corresponding to the fundamental frequency of the speech stimulus (103– 121 Hz), first formant frequencies of the stimulus (454- 719 Hz) and for the higher harmonics (721-1155 Hz) were measured for all the subjects.

CHAPTER 4

RESULTS AND DISCUSSION

The present study was an effort to compare the auditory plasticity in non-musicians & musicians and to find out correlation between brainstem encoding & perception of speech in noise. The brainstem responses were measured by speech evoked auditory brainstem response in quiet condition. Speech evoked LLR was obtained to measure cortical responses (P1, N1, and P2). Speech Perception in Noise (SPIN) test was done at 0 dB SNR separately for both the ears. A total of 25 non-musicians and 25 trained Carnatic vocal musicians participated in the study, who were classified in to 3 groups based on their musical experience or training. The data was appropriately tabulated and statistically analyzed using SPSS (version 18) software.

The Following analyses were carried out:

1. Descriptive statistics (mean and standard deviation) were obtained for all the parameters for both ears separately.

2. Separate 2-WAY MANOVA was done to see the significant difference between musician & non-musicians for all the parameter for speech ABR (Latency of wave V,A ,C ,D , E ,F ,O), LLR (Latency of P1 ,N1 ,P2) and for amplitudes of Fo & F1.
3. 2-WAY ANOVA was done to see the significant difference between musicians & non-musicians for SPIN scores.
4. Pearson correlation was calculated to see the correlation between SPIN scores and the amplitudes of Fo & F1 in musicians & non-musicians.

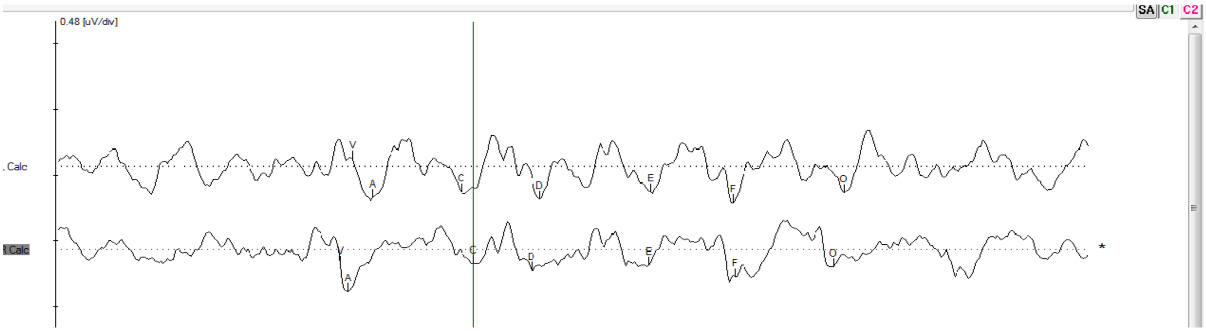

Figure 4.1 Speech-Evoked ABR of an individual of Group

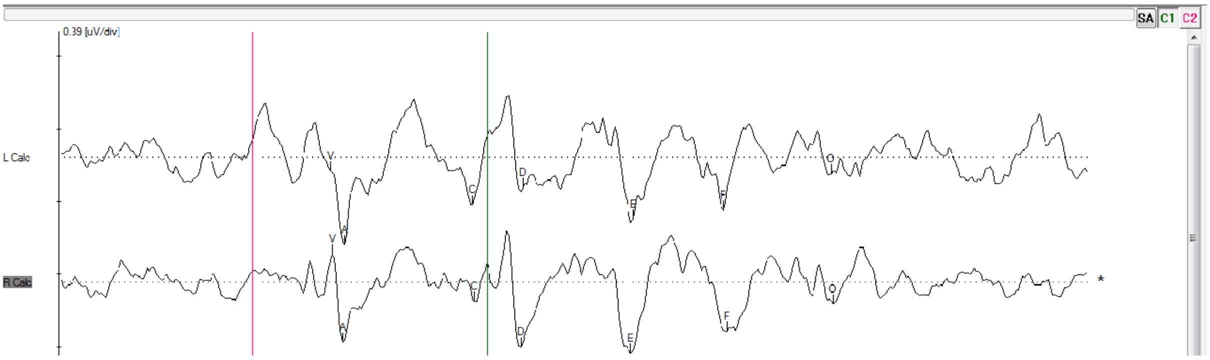

Figure 4.2 Speech-Evoked ABR of an individual of Group 2

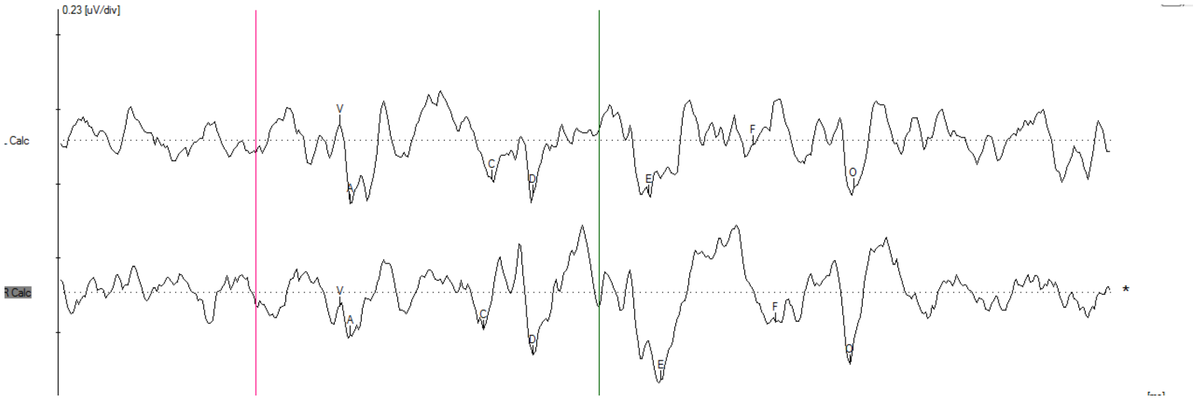

Figure 4.3 Speech-Evoked ABR of an individual of Group 3

SPEECH ABR

Table 4.1*Mean and standard deviation of waves of speech evoked ABR in non-musicians*

Waves	Age Group	Mean(ms)	SD
V	1	6.9960	.15565
	2	7.0325	.21008
	3	7.1120	.26355
A	1	7.9540	.31124
	2	7.7570	.25676
	3	7.8910	.24333
C	1	18.4060	.22016
	2	18.4570	.22841

	3	18.4825	.20865
D	1	22.7130	.28194
	2	23.5250	2.18359
	3	23.2255	.46778
E	1	31.8600	.83775
	2	31.5840	.24040
	3	31.5850	.54375
F	1	39.5450	.33870
	2	39.2285	1.10124
	3	39.5250	.24239
0	1	48.1200	.26179
	2	48.2830	.50602
	3	48.0405	.32528

Table 4.2*Mean and standard deviation of waves of speech evoked ABR* in musicians

Wave	Age Group	Mean(ms)	SD
V	1	7.4170	.39169
	2	6.9750	.15979
	3	6.9450	.10899
A	1	8.0990	.13093
	2	7.7770	.22483
	3	7.7990	.22592
C	1	18.5050	.30226
	2	18.4370	.20989
	3	18.5035	.26230
D	1	22.7160	.17896
	2	23.2505	.37410
	3	22.7695	.35933
E	1	31.5350	.20587
	2	31.5205	.29790
	3	31.6355	.39321
F	1	39.5820	.33072
	2	38.9255	1.89959
	3	39.5300	.30252
0	1	48.2560	.40533
	2	47.9065	1.00969
	3	48.2960	.47158

Latency of Waves V, A, C, D, E, F & O was measured for non-musicians & musicians, for both ears separately, for all three groups. Table 4.1 and 4.2 shows descriptive statistics (mean & SD) for latencies of speech evoked ABR waves for musicians & non-musicians across three groups.

2-way MANOVA was done to see the differences between musicians and non-musicians for latencies of speech evoked ABR (TABLE 4.3).

TABLE 4.3 *Statistical values*

Source	Variable	df	F	Sig.
GrpNM*M	V	2	12.596	.000*
	A	2	1.707	.187
	C	2	.425	.655
	D	2	.330	.720
	E	2	1.227	.298
	F	2	.298	.743
	O	2	3.105	.049*

*The mean difference is significant at the .05 level.

There was a significant difference between non-musicians & musicians ($p<0.05$) for latency of wave V & O for all three groups. There was no significant difference present for latencies of other waves between musicians and non-musicians.

In the present study, the latency of wave V and O responses were significantly different between musicians and non-musicians. There was no significant difference present in transition latencies between musicians and non-musicians. These results are in agreement with the study by Parbery-Clark et al (2009), where it was concluded that musicians had earlier response onset timing, than non-musicians. Musacchia et al (2008) reported that latency and amplitude of wave V differed between musicians and non-musicians. The results of the study by Parbery-Clark et al (2009) also suggested that there was no significant difference in transition latencies between musician and non-musician in quiet conditions. Musicians and non-musicians had equivalent stimulus response correlation in quiet. Their results are in agreement with the results of the present study.

SPEECH EVOKED LATE LATENCY RESPONSE

TABLE 4.4 *Mean and standard deviation (SD) of waves of speech evoked LLR in non-musician.*

Wave	Age Group	Mean(ms)	SD
P1	1	90.3290	8.78046
	2	88.9465	8.41858
	3	89.1212	11.56455
N1	1	143.1100	13.68969
	2	188.0965	225.18617
	3	191.7613	219.84187
P2	1	184.2470	14.84895
	2	175.1400	14.91415
	3	186.0255	17.74305

TABLE 4.5 *Mean and standard deviation (SD) of waves of speech evoked LL*R in musician

Wave	Age Group	Mean(ms)	SD
P1	1	91.1380	8.53371
	2	82.7090	13.75644
	3	82.6675	14.56331
N1	1	146.9080	10.76211
	2	136.4030	11.75320
	3	129.2270	14.66725
P2	1	177.2520	13.80140
	2	182.3605	14.68192
	3	182.1315	20.68962

Latency of Waves P1, N1& P2 was measured for non-musicians & musicians, for both ears separately. Table 4.4 & Table 4.5 show descriptive statistics (mean & SD) for latencies of speech evoked LLR waves across all age groups.

2-way MANOVA was done to see the differences between musicians and non-musicians for latencies of speech evoked LLR waves across years of musical experience (TABLE 4.6)

TABLE 4.6 *Statistical values*

Source	Variable	df	F	Sig.
grpNM*M	P1	2	.749	.476
	N1	2	.383	.683
	P2	2	1.650	.198

*The mean difference is significant at the .05 level.

TABLE 4.6 reveals that there is no significant difference between musicians & non-musicians in terms of latencies of wave P1, N1 and P2 of speech evoked LLR ($p < 0.05$).

The results of the present study are in consonance with the results by Strait et al (2011). There was no response variability among musicians and non-musicians at any electrode site.

Shahin et al (2003) reported enhanced P2 and N1c responses in musicians compared to non-musicians. Krista et al (2009) reported that long term music training offers structural plasticity in developing correlation with behavioural changes. T1 weighted MRI was used in their study for assessment. The difference in the results of the present study with the earlier studies reported in the literature can be accounted on the following reason: First, the assessing tool used in previous studies is magneto encephalography (MEG), electroencephalography (EEG) and MRI. These radiological tests are different from far field electrophysiological responses. Second, the previous studies were conducted on instrumental musicians, whereas the present study was carried out on Carnatic vocal musicians. Moreover, the subjects taken in Shahin et al (2003) study were having more years of musical experience (greater than 11 years) than the subjects of the present study.

FFT- Fast Fourier Transform; F0, F1 & F2.

TABLE 4.7 *Mean and SD for FFT- Fast Fourier Transform; Fo, F1 & F2- Fundamental frequency, first and second Formants for non-musicians.*

Wave	Age Group	Mean(ms)	SD
Fo	1	2.07690	.645367
	2	3.37990	1.308120
	3	4.06230	1.204132
F1	1	1.41260	.368791
	2	1.28670	.343958
	3	1.25225	.434901
F2	1	.53920	.095237
	2	.44760	.115118
	3	.46075	.162216

TABLE 4.8 *Mean and SD for FFT- Fast Fourier Transform; F0, F1 & F2- Fundamental frequency, first and second formants for musicians.*

Wave	Age Group	Mean(ms)	SD
Fo	1	2.30340	.669302
	2	4.97560	.971714
	3	5.55440	1.196472
F1	1	1.32690	.556053
	2	1.51805	.475621
	3	1.38995	.381008
F2	1	.48280	.153182
	2	.52360	.136830
	3	.52705	.149395

Fo, F1 & F2- Fundamental frequency, first and second formants were measured for non-musicians & musicians, for both ears separately, for all three groups. Table 4.6 and Table 4.7 show descriptive statistics (mean & SD) for FFT- Fast Fourier Transform; Fo, F1 & F2- Fundamental frequency, first and second formants for non-musicians & musicians.

2-way MANOVA was done to see the differences between musicians and non-musicians for FFT- Fast Fourier Transform; F0, F1 & F2- Fundamental frequency, first and second Formants (TABLE 4.9).

TABLE 4.9 *Statistical values*

Source	Variable	df	F	Sig.
grpNM*M	Fo	2	2.910	.000*
	F1	2	.934	.397
	F2	2	1.689	.190

*The mean difference is significant at the .05 level.

There was a significant difference present ($p< 0.05$) for Fo Formant between musicians and non-musicians for group 3. The amplitude of energy concentration in Fo formant is significantly larger in Group 3 compared to non-musician. The results of the present study are in agreement with results of Mussachia (2007). He reported that musicians have larger response amplitudes for encoding of speech and music stimuli compared to non-musicians. Mussachia et al (2008) reported experienced musicians had larger Fo peak amplitudes. In the present study, musicians in Group 3, with higher years of musical experience had better mean SPIN scores than non-musicians. This result draws support study by Anderson et al (2010), which suggests that good SIN perceivers had greater spectral magnitudes for Fo and H2.

The difference in the results for groups 1 and 2 of the present study can be explained on the following reasons; first, the musical year of experience in group 1 and 2 were less than reported in the previous studies. Second, the previous studies were conducted on instrumental musicians, whereas the present study was carried out on vocal musicians. Moreover, the subjects taken in Mussachia et al (2008) study were having more experience than the subjects in the present study. Most of the study reports experience of 10 for their subjects.

Speech Perception in Noise

The speech perception in noise was assessed for all the 50 subjects for both the ears. The test was carried out at 0 dB SNR.

TABLE 4.10 *Mean and SD for SPIN for all groups.*

Group	NM/M	Mean (%)	SD
1	Non musician	72.80	3.676
	Musician	73.20	4.237
2	Non musician	77.00	4.657
	Musician	80.00	4.768
3	Non musician	80.80	3.488
	Musician	86.00	2.865

2-WAY ANOVA was employed to see the significant difference between musicians & non-musicians for SPIN scores (TABLE 4.11).

TABLE 4.11 *Statistical values*

Source	df	F	Sig.
GROUP	2	32.512	.000*
NM_M	1	3.262	.074
GROUP * NM_M	2	1.061	.350

There was a significant difference across the three groups *(p*< 0.05). However, there was no significant difference between musicians & non-musicians across the different groups for SPIN.

On comparing mean score (TABLE 4.10) the mean scores are quite better For musicians in group 3, but not significantly better. But this is in contrast to the previous research done on speech perception abilities in musicians. According to a study done by Parbery-Clark et al (2009), musical experience enhances the ability to hear speech in challenging listening environments. In another study Parbery-Clark et al (2009) found that musical experience resulted in more robust subcortical representation of speech in the presence of background noise. The difference in the results of the present study with the earlier studies reported in the literature can be accounted on the following reasons: First, the noise used in the previous studies were speech shaped noise or multi-talker babble. But in the present study speech noise was used to study the speech perception in noise. It is evident that the speech shaped noise or multi-talker babble will give better results for speech perception in noise when compared to speech noise. Second, the previous studies were conducted on instrumental musicians, whereas the present study was carried out on vocal musicians. Moreover, the subjects taken in Parbery-Clark et al (2009) study were having more experience than the subjects in the present study. Third, the speech material used in previous studies was sentences (Quick SIN, HINT). The sentences are more redundant than words. Fourth, the present study was conducted at 0 dB SNR, the studies reported in literature suggested that SPIN is better in adverse listening conditions. Parbery-Clark (2009) reported musicians were able to repeat sentences presented at a lower, more challenging SNR than non-musicians.

CORRELATION BETWEEN SPIN & FFT

Correlation between SPIN and FFT were (TABLE 4.12).

TABLE 4.12 *Correlation between FFT and Speech perception in noise.*

Formants	SPIN
Fo	.000**
F1	.826
F2	.914

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

The results revealed that there is a positive and highly significant correlation between SPIN & Fo. With the increase in Fo amplitude, there is an increase in the scores for the perception of speech in noise. This finding can be supported by study of Anderson et al (2010), good SIN perceivers have greater spectral magnitudes for Fo and H2.

There was no significant difference between other formants & SPIN scores. As the Formants increases, the amplitude of the harmonics decreases.

Thus the results of the present study reveal that latencies of wave V and O were significantly different between musicians and nonmusicians.

The latencies of speech evoked LLR was does not reveal statistically significant difference for musician and nonmusician. A significant difference was noticed between musician and nonmusician for Fo formant. There is a significant difference across the three groups of musicians on SPIN score. A highly significant positive correlation is present between Fo and SPIN scores.

CHAPTER 5

SUMMARY & CONCLUSION

The present study was aimed to find out the Brainstem correlation of speech in noise perception in musicians and to compare auditory plasticity in musicians and non-musicians. A total of 50 Carnatic vocal musicians participated in the study. The musicians were classified into three groups based on their age and experience. Auditory plasticity was measured by using speech evoked ABR and LLR. Fast Fourier Transform was done to find out energy concentration in formants. Speech perception in noise was measured at 0 dB SNR.

Following conclusions can be drawn from the results of the present study:

1. Fo encoding is better in musicians than in non musicians.
2. Speech perception ability in musicians becomes better with increased years of musical exposure and experience.
3. The increase in Fo amplitude is positively correlated with the speech perception in noise.
4. There was no significant difference in latencies of P1, N1 and N2 between musicians and non musicians for speech evoked LLR.

Implications

- To add information to the literature.
- Can be implemented in Hearing Aid technology for musicians with hearing loss to improve their speech perception.

- Music training can be used as a potential remediation strategy for children requiring language training and auditory processing disorders with noise exclusion deficits.

- Future research on clinical population who may exhibit neural encoding deficits such as autism. Brainstem maturation as an indication in infants & preschool children at risk.

Future Directions for Research

- The present study can b replicated across vocal musicians and instrumental musicians.
- Can be compared between Hindustani and Carnatic musicians,
- Musicians and dancers can be compared to find whether there are differences in Fo encoding and ability to perceive speech in the presence of noise.

REFERENCES

Anderson, S., Skoe, E., Chandrasekaran, B,. Zecker, S., & Kraus, N,. (2011) Brainstem Correlates of Speech-in-Noise Perception in Children. *Hearing Research* ,270,151-157.

Anderson, S., Parbery-Clark, A., Yi , H., & Kraus., N., (2011). A neural Basis of speech-in-noise perception in older adults. *Ear & Hearing*, 32, 1-8.

Anderson, S., Kraus., N., (2010). Sensory-cognitive interaction in the neural encoding of speech in noise: A review (2010). Journal of American Academy of Audiology 21, 575-585.

Assmann, P. F., & Summerfield, Q. (1990). Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies. *Journal of Acoustical Society of America, 88*, 680-697.

Besson, M., Schon, D., Moreno, S., Santos, A., & Magne, C. (2007). Influence of musical expertise and musical training on pitch processing in music and language. *Restorative Neurology and Neuroscience, 25*, 399-410.

Bialystok, M.W., & DePape, A. M. (2009). Musical expertise, bilingualism, and executive functioning. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 565-574.

Beavois, M. W., & Meddis, R. (1997). Time decay of auditory stream biasing. *Perception and Psychophysics, 59*, 81-86.

Brokx, J. P. L., Nooteboom, S. G., & Cohen, A. (1982). Intonation and the perceptual separation of simultaneous voice. *Journal of Phonetics, 10*, 23-36.

Chan, A. S., Ho, Y. C., & Cheung, M. C. (1998). Music training improves verbal memory. *Nature, 396*, 128.

Chandrasekaran., B., Kraus., N., (2010). The scalp-recorded brainstem response to speech : Neural origins and plasticity. *Psychophysiology*. 47, 236–246.

Chandrasekaran., B., Kraus., N., (2010). Music, Noise- Exclusion, and learning. *Music perception* 27, 297-306.

Elbert, T., Pantev, C., Wienbruch, C., Rockstroh, B., Taub, E. (1995). Increased cortical representation of the fingers of the left hand in string players. *Science, 270(5234),* 305-307.

Foregeard, M., Winner, E., Norton, A., & Schlaug, G. (2008). Practicing a musical instrument in childhood is associated with enhanced verbal ability and nonverbal reasoning. *PLoS One*, *3*, 3566.

Fujioka, T., Trainor , L., Ross, B., Kakigi , R., Musical Training Enhances Automatic encoding of Melodic Contour and Interval Structure(2004*). Journal of cognitive Neuroscience* 16, 1010-1021.

Gaser, C., Schlaug, G. (2003). Brain structures differ between musicians and non- musicians. *Journal of Neurosciences*, *23 (27)*, 9240-9245.

Hebb, D. O., (1949). The organization of behavior. New York: Wiley.

Hoormann, J., Falkenstein, M., Hohnsbein, J., & Blanke, L. (1992). The human frequency following response (FFR): normal variability and relation to the click-evoked brainstem response. *Hearing Research, 59 (2),* 179-188.

Johnson KL, Nicol T, Kraus N. (2008) Developmental plasticity in the human auditory brainstem *Journal of Neuroscience* 28(15): 4000-4007.

Kraus., N., Skoe, E., Parbery-Clark, A., & Ashley, R. (2009). Experience induced malleability in neural encoding of pitch, timbre and timing: implications for language and music. *Annals of New York Academy of Sciences* 1169, 543-557.

Kraus, N., & Banai, K. (2007). Auditory-processing malleability: Focus on language and music. *Current Directory of Psychological Sciences, 16,* 105-110.

Krishnan, A., Xu, Y., Gandour, J., Cariani, P. (2005). Encoding of pitch in the human brainstem sensitive to language experience. *Brain Research Cognitive Brain Research, 25 (1),* 161-168.

Margulis, E. H., Mlsna, L.M., Uppunda, A. K., Parrish, T. B., Wong, P. C. (2009). Selective neurophysiologic responses to music in instrumentalists with different listening biographies. *Human Brain Mapping, 30*, 267- 275.

Menning, H., Roberts, L., Pantev, C., (2000). Plastic changes in the auditory cortex induced by intensive frequency discrimination training. *Auditory and Vestibular Systems* 11, 817- 822.

Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences of United States of America, 104,* 15894-15898.

Nair, D., Large, W. E., Steinberg, F., & Kelso, J. A. S. (2002). Expressive timing and perception of emotion in music: an fMRI study. *Proceeding of the 7th International Conference on Music Perception and Cognition, 627*, 627-630.

Ohinshi, T., Matsuda, H., Asada, T., Aruga, M., Hirakata, M., Nishikawa, M. (2001). *Cerebral Cortex, 11(8)*, 754-760.

Pantev, C., Roberts, L. E., Schulz, M., Engelien, A., Almut., Ross.,& Bernhard. (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport*, *12*, 169–174.

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musical experience limits the degradative effects of background noise on the neural processing of sound. *Journal of Neuroscience*, *29,* 14100-14107.

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech in noise. *Ear and Hearing. 30*, 653-661.

Russo N, Nicol T, Zecker S, Hayes E, Kraus N. (2005) Auditory training improves neural timing in the human brainstem. *Behavioural Brain Research* 156: 95-103.

Schalaug, G. (2001). The brain of musicians: a model for functional and structural adaptation. *Annals of New York Academy of Sciences*, *930*, 281-299.

Schneider, P., Scherg, M., Dosch, H. G., et al. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Natural Neuroscience, 5,* 688-694.

Strait, D., Kraus, N., Parbery-Clark, A., & Ashley, R. (2010). Musical experience shapes top-down auditory mechanisms: Evidence from masking and auditory attention performance. *Hearing Research, 261*, 22-29.

Strait, D., Kraus, N., Skoe, E., & Ashley, R. (2009). Musical experience and neural efficiency: Effects of training on subcortical processing of vocal expressions of emotion. *European Journal of Neuroscience, 29,* 661-668.

Sonitha (2011) Unpublished dissertation AIISH

Shahin , A., Bosnyak , D., Trainor, L., Roberts , Larrey, R., (2003).Enhancement of neurolplastic P2 and N1c auditory evoked potentials in musicians. The Journal of Neuroscience, 12 ,5545–5552 .

Song , H., Skoe, E., Patrick C., Kraus, N.,(2010).*Journal of Cognitive Neuroscience* 10, 1892-1902.

Suga, N., Zhang, Y., Yan, J. (1997). Sharpening of frequency tuning by inhibition in the thalamic auditory nucleus of the mustached bat. *Journal of Neurophysiology, 77*, 2098-2114.

Thomas, A., (2011). Unpublished dissertation AIISH

Tremblay, K., Kraus, N., McGee, T., Ponton, C., & Otis B. (2001). Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. *Ear & Hearing*, 22, 79–90.

Tzounopoulus, T., Kraus, N., Learning to encode timing : Mechanisms of plasticity in auditory brainstem.*Neuron* 62, 463-469.

Wong, P. C., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English speaking adults. *Applied Psycholinguistics, 28,* 565-585.

Wong, P. C., Skoe, E., Russo, N. M., Dees, T., & Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience, 10*, 420-422.

Yathiraj, A., & Vijayalakshmi, C. S. (2005). *Phonemically Balanced word list in Kannada*. Developed in Department of Audiology, AIISH, Mysore.

Zendel, B. R., Alain, C. (2009). Concurrent sound segregation is enhanced in musicians. *Journal of Cognitive Neuroscience, 21*, 1488-1498.

Zhang, Y., Suga, N., Yan, J. (1997). Corticofugal modulation of frequency processing in bat auditory system, *Nature, 387*, 900-903.