Comparison of temporal resolution abilities and speech perception in noise in children born in families with musical background and children born in families without musical background

CHAPTER 1

INTRODUCTION

"Without music, life is a journey through a desert."

- *Pat Conroy*

Music can be ubiquitous in every culture all around the world. Earlier researchers have focused on how formal music training impacts various aspects of cognitive development such as auditory perception, memory, and language skills. The present study was aimed at comparing the temporal resolution abilities and speech perception in noise for children born in families with and without musical background. The perception of music involves complex brain functions underlying acoustic analysis, auditory memory, auditory scene analysis and processing of musical syntax. Moreover, music perception potentially affects emotion, influences autonomic nervous system, the hormonal and immune systems and activates (pre) motor representations.

Many studies have reported that musicians have better auditory perception skills when compared to non-musicians. There are many studies in literature which have documented that musical training improves basic auditory perceptual skills resulting in enhanced behavioral (Jeon & Fricke, 1997; Koelsch et al., 1999; Oxenham et al., 2003; Tervaniemi et al., 2005; Micheyl et al., 2006; Rammsayer & Altenmuller, 2006) and neurophysiological responses (Brattico et al., 2001; Pantev et al., 2001; Schneider et al., 2002; Shahin et al., 2003, 2007; Tervaniemi et al., 2005; Kuriki et al., 2006; Kraus et al., 2009). Musicians' life long experience in detecting melodies from background harmonies can be considered as a process analogous to speech perception in noise. Research reports that musicians had a more robust sub- cortical representation of the acoustic stimulus in the presence of noise (Kraus et al., 2009). Musical practice not only enhances the processing of music related sounds but also influences processing of language (Marques et al., 2007; Moreno et al., 2009; Parbery-Clark et al., 2009a; Schon et al., 2004). Because of their musical training, musicians have learned to pay more attention to the acoustic details of the stimulus than non-musicians.

Musical training involves discrimination of pitch intonation, onset, offset and duration aspects of sound timing as well as the integration of multisensory cues to perceive and produce notes. Because of their musical training, musicians have learned to pay more attention to the details of the acoustic stimuli than non-musicians (Musacchia et al., 2007). Music is a complex auditory task and musicians spend years for 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).

Musicians spend most of their time attending to and manipulating complex auditory signals that comprise multiple streams. In addition to processing concurrent auditory units (i.e., simultaneously occurring melodies), musicians must also analyze the vertical relationships between streams (i.e., Harmony). In addition to this online auditory scene analysis musicians also hone their abilities to conceive, plan, and perform music in real time. Previous work has documented that musical training improves basic auditory perceptual skills resulting in enhanced behavioral (Jeon & Fricke, 1997; Koelsch et al, 1999; Oxenham et al. 2003; Tervaniemi et al. 2005; Rammsayer & Altemuller 2006) and neurophysiological responses (Brattico et al. 2001; Pantev et al. 2001; Schneider et al. 2002; Tervaniemi et al. 2005; Kraus et al. 2009). Moreover, it would seem that musicians are able to use these perceptual benefits to facilitate concurrent sound segregation (Zendel & Alain 2009). Musical training not only enhances aspects that are specific to musical perception, but these enhancements also cross over to other domains, particularly language, suggesting shared neural resources for language and music processing (Patel 2003, 2007; Kraus & Banai 2007; Kolesch et al. 2008). For example, lifelong musical experience is linked to improved subcortical and cortical representations of acoustic features important for speech encoding and vocal communication (Magne et al. 2006; Schon et al. 2004; Marques et al. 2007; Musacchia et al. 2007, 2008; Chandrashekaran et al. 2008; Strait et al. 2009). Likewise, musical experience has been shown to improve verbal ability (Forgeard et al. 2008), verbal working memory and verbal recall (Chan et al. 1998; Brandler & Rammsayer 2003; Ho et al. 2003; Jackobsen et al. 2003). As a consequence of the musician's extensive experience with auditory stream analysis within the context of music, more honed auditory perceptual skills as well as greater working memory capacity, musicians seem well equipped to cope with the demands of adverse listening situations such as Speech in Noise.

Previous studies also report that music training benefits auditory processing not only in the musical domain, but also in the processing of speech stimuli (Musacchia et al., 2007; Schon, Magne & Besson, 2004; Wong, Skoe, Russo, Dees & Kraus, 2007). Consistent findings across a range of studies that use methods spanning from neurophysiology to behavior indicate that music training improves a variety of verbal and non verbal skills. These include working memory (Chan, Ho & Cheung, 1998; Forgeard, Winner, Norton & Schlaug, 2008), processing of prosody and linguistic features in speech (Chandrasekaran, Krishanan & Gandour, 2009; Wong et al., 2007), phonological skills (Forgeard, et al. 2008), processing emotion in speech (Strait, Kraus, Skoe & Ashley, 2009), auditory attention (Strait, Kraus, Parbery- Clark & Ashley, 2010) and auditory Stream segregation (Beauvois & Meddis, 1997). The above findings suggest that there are many benefits of music training; one among them is the auditory processing. Temporal resolution is one of the domains of auditory processing

Temporal Resolution is defined as the perception of a short interval of time that each individual can discriminate between two auditory signals which is of about 2-3 ms. The studies suggest that exposure to sound during the first two years of life is important for the maturation of the structures of the central nervous system. The contact with the music before the age of seven could contribute to the development of the primary auditory cortex more precisely the planum temporale. The musicians surveyed had an increase in the left temporal plane identified by investigations of magneto encephalography. The different studies on musicians suggest that musical training diary, used by professional musicians, can induce functional reorganization of the cerebral cortex. Musicians have better neural activation due to long term musical training.

Previous research points out that, better development of planum temporale is seen in those individuals where the musical training had begun before the age of nine. However, other studies argue that musical ability is innate and that musical training is not responsible for the development of planum temporale. However, it is confirmed that there is development of the planum temporale in relation to individuals who were exposed to early musical stimulus (Pantev et al., 2001). Another study by Ishll et al. (2006) shows that music has a positive influence on the development of the planum temporale, because according to their study, subjects exposed to musical training (singing) for over four years compared to amateur musicians without professional guidance, performed better on temporal resolution through the test Random Gap Detection Threshold (RGDT).

Speech perception in noise (SIN) is a complex task requiring the segregation of the target signal from the competing background noise. This task is further complicated by the degradation of the acoustic signal, with the noise particularly disrupting the perception of the fast spectro-temporal changes (Brandt & Rosen, 1980). Whereas, children with language-based learning disabilities (Bradlow et al., 2003; Ziegler et al., 2005) and hearing impaired adults (Gordon- Salant & Fitzgibbons, 2005) are especially susceptible to the negative effects of background noise, musicians are less affected and demonstrate better performance on SIN when compared to non-musicians (Parbery-Clark et al., 2009). Compared to non-musicians, musicians exhibit enhanced subcortical encoding of sounds with both faster responses and greater frequency encoding. These enhancements are not simple gain effects. Rather, musical experience selectively strengthens the underlying neural representation of sounds reflecting the interaction between cognitive and sensory factors (Kraus et al., 2009), with musicians demonstrating better encoding of complex stimuli (Wong et al., 2007) as well as behaviorally relevant acoustic features (Lee et al., 2009).

In order to extract the target acoustic signal, our auditory system must resolve two issues. First, there must be a process that partitions the acoustic input into separate auditory units. Second, there must be a mechanism for appropriately organizing these acoustic units over time. Auditory scene analysis is the term given to the internal process of segregating and subsequent grouping of an auditory stream (Bregman, 1990). Auditory scene analysis is based on the notion that pre-attentive processes use the Gestalt laws of organization (Koffka, 1935) - physical similarity, temporal proximity, and good continuity- to group sounds. In acoustic terms, sounds with similar frequency and spatial location are more likely to be grouped together as auditory units. Indeed listeners take advantage of both frequency and spatial location cues to assist in the perception of SIN. Perceptual streaming, or the ability to hear two streams, is facilitated when concurrently presented complex tones are separated by as little as one semitone. For example, when asked to identify simultaneously presented vowels, performance improved when the fundamental frequencies were different (Scheffers, 1983; Assmann & Summefield, 1990). This phenomenon can help to explain why speech recognition in noise is more difficult when the target and the background speakers are of the same sex, and the fundamental frequencies of different voices are consequently closer in frequency. Even small frequency differences between speakers' voice can be used as cues to aid speaker differentiation (Treisman, 1964; Brkox & Nooteboom, 1982).

The ability to properly group, represent, and store auditory units over time is fundamental to forming auditory streams and is therefore an essential aspect of SIN perception. Concurrently presented auditory units may be represented as separate, parallel sensory traces that are not completely independent of each other (Fujioka et al., 2005, 2008). This not only highlights the auditory system's ability to represent simultaneously presented auditory units as both separate yet integrated sensory traces (Fujioka et al., 2005,2008) but also support the idea that stream segregation is an active, rather than a passive process (Alain & Brenstein 2008).

Need for the Study:

The studies mentioned in the introduction section have highlighted the better auditory perceptual skills in trained musicians when compared to non-musicians. But there are no studies to demonstrate the temporal resolution and speech perception abilities in children born in families with musical background. Hence, it is interesting to study the effects of familial musical background influencing the auditory abilities in children.

Aim of the Study:

The present study aimed to:

- Find the temporal resolution abilities of children born in families with musical background and children born in families with no musical background.
- Find the speech perception abilities in presence of background noise for the same group.

CHAPTER 2

REVIEW OF LITERATURE

The following section provides a brief review of literature regarding the effect of musical training and familial influence of music on children in terms of temporal resolution and speech perception ability in the presence of background noise. The review has been divided into five sections mainly: 1) Effect of music training on structural functional changes in the nervous system. 2) Effect of music training on language related skills, 3) Effect of music training on emotional and cognitive processes, 4) Effect of music training on perception in noise and 5) Effect of music training on temporal abilities.

Effect of music training on structural and functional changes in the 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 & Schlaug, 2003) and have more axonal projections that connect the right and left hemispheres (Schlaug et al., 2005). 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 (Ohinshi 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). With regard to evoked potentials which 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).

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 showed 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 was that recent musical training improves one's auditory memory and shapes composite (P1-N1) and pitch encoding (F0) in a co-coordinated manner. 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.

Several studies show differences between the brain of adult musicians and non-musicians. For example, structural MRI studies indicate differences in grey 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 & directionality 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.

At a functional level, the brain responses of adult musicians and non-musicians also differ as measured by EEG and MEG. 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). For 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 is present in both musicians and non musicians, it is 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). Errors in one chord of a chord sequence produce an early right anterior negativity that is also larger in musicians than in non-musicians (Koelsch et al., 2002).

Effect of Music Training on 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 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. Further, in tonal languages, changes in pitch are used lexically; that is, in differentiating between words. 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 (Chandrasekaran, 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). 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. Musicians showed superior cortical representation of linguistic pitch in a non-native language relative to non musicians (Chandrasekaran et al., 2009). In this 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, 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). 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, which 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 report 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 provided 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). Similarly musicians also demonstrated selective neural enhancement of the upper note of musical chords (Lee, Skoe, Kraus & Ashley, 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; 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).

Effect of music training on perception 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 sub serve successful perception of speech in noise. A recent 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 practice fine tunes cognitive and sensory abilities, leading to an overall advantage for 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.

Effect of music training on temporal abilities

Monteiro et al. (2010) compared the temporal resolution abilities in musicians and non-musicians. The study was characterized by prospective and compared between two groups, one consisting of 20 musicians and other 20 non musicians matched for age and education underwent audiological evaluation and Gap in Noise test (GIN) to evaluate the temporal resolution. The test performance of the GIN group of musicians was not significant in the control group in the right ear (RE) or left (LE). The correlation between the average high frequencies for the LE with the GIN test was (p= 0.001) in the control group. The average frequencies for both ears in the group of musicians was statistically significant and the highest values for RE (p= 0.0001). There was no difference between the performances of the GIN test for both groups as well as the correlation between duration of daily exposure to music and GIN.

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. Hence, it would be curious to observe the influence of musical families on their children in terms of their auditory abilities.

CHAPTER 3

METHOD

The present study aimed to compare the temporal resolution abilities and speech perception in noise in children born in families with musical background and children born in families without musical background for which, the following methodology was undertaken.

Participants:

The participants were classified into 2 groups.

Control Group:

Subgroup A: 10 children born in non-musical background with no formal musical training in the age range of 12-16 years were considered. This group is denoted as ‘S_A’.

Subgroup B: 10 Children born in non-musical background but trained in music, in the age range of 12-16 years were considered. This group is denoted as ‘S_B’.

Experimental Group:

Subgroup 1: 10 age matched children born with musical background but not trained in music were considered. This group is denoted as ‘S_1’.

Subgroup 2: 10 age matched children born with musical background and also undergoing formal musical training were considered. This group is denoted as ‘S_2’.

An informal interview was administered to all the participants in order to obtain the information about musical background.

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 226 Hz probe tone with normal acoustic reflexes in both ears.
- Speech recognition Threshold of ±12 dB (re. PTA of 500, 1000 & 2000 Hz) in both ears.
- Speech identification scores of > 90% at 40 dB SL (re. SRT) in both ears.
- No indication of rectrocochlear pathology (RCP).
- No history of neurological or otological problems
- No illness on the day of testing.
- There was a clear bifurcation of subjects with respect to the musical background.

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 TDH-39) housed with MX-41/AR ear cushions with audio cups and a bone vibrator (Radio ear B71) were used to assess the pure tone threshold.
2. GSI Tympstar (Grason-Stadler Inc. USA) middle ear analyser was used for tympanometry and reflexometry.
3. A laptop (Acer Aspire 5050) was used to deliver the stimulus for Gap Detection Test (GDT), Temporal Modulation Transfer Function (TMTF) and Quick Speech in Noise (QuickSIN) - Kannada.

Stimuli

- Stimuli for GDT & TMTF were generated through psychoacoustic toolbox module run under Matlab R2010b software.
- Quick Speech in Noise test in Kannada developed by Avinash, Methi., & Kumar, 2009).

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, 1976) were used to obtain the speech Recognition Threshold (SRT) from both the ears. A set of 3 spondees were presented at 20 dB SL 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.

Immittance Audiometry

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

Temporal Modulation Transfer Function (TMTF)

Two stimuli, unmodulated white noise and sinusoidally amplitude modulated (SAM) white noise of 500 ms duration, with a ramp of 20 ms were used. The stimuli was generated using a 32 bit digital to analog converter with a sampling frequency of 44.1 KHz and were low pass filtered with a cut-off frequency of 220 Hz. The Modulated signal was derived by multiplying the white noise by a dc-shifted sine wave. The depth of modulation was controlled by varying the amplitude of modulating sine wave. Modulation depth was varied between 0 to -30 dB (where 0 dB is equal to 100% modulation depth and -30 dB is equal to 0% modulation). Six different modulation frequencies were used (4 Hz, 8 Hz, 16 Hz, 32 Hz, 64 Hz, & 128 Hz).

The stimuli were presented at 40 dB SL (with reference to PTA) or at comfortable level. The stimuli were presented to the participants through headphones. The subjects' task was to discriminate between modulated and the unmodulated noise till they were able to identify the difference.

Three interval alternate forced choice methods (3IAFC) were used. On each trial, un-modulated and modulated stimuli were successively presented with an inter-stimulus interval of 500 ms. Modulation depth was converted into decibels [20 log 10(m), where 'm' refers to the depth of modulation]. A step size of 4 dB was used initially and then reduced to 2 dB after two reversals. This procedure provides an estimate of the value of amplitude modulation necessary for 70.7 % estimate of correct responses (Levitt, 1971). The mean of eight reversals in a block of 14 was taken as threshold.

Gap Detection Threshold (GDT)

Gap detection test consisted of a standard stimulus of 750 ms duration Gaussian noise with a silence of standard duration placed at its temporal center. The variable stimulus had variable gap duration and the length of its gap was changed as a function of the subject's performance. All noises had a 0.5 ms cosine ramp at both onset and offset. Three Interval Alternate Forced Choice Method (3IAFC) method was used to obtain the gap detection threshold. It consisted of three blocks of white noise, one of which contained gaps of variable duration. The subjects' task was to identify the gap and to detect which block of noise was having the gap in it. The presentation level of the stimulus was 40 dB SL (with reference to PTA) or most comfortable level, monaurally. Each time when the subject detected the gap embedded in noise correctly, the size of the gap was reduced and test was continued till the subject could trace the smallest gap. The minimum gap that the subject detected was considered as the gap detection threshold. The gap detection thresholds were obtained for both the groups.

Quick Speech Perception in Noise - Kannada

Quick Speech Perception in Noise - Kannada (Avinash, Methi & Kumar, 2009) was done using the 60 sentences based on the subjects rating of predictability. 60 sentences which were distributed in 12 lists with 7 sentences in each list were used. Sentences were randomly divided into different lists. Some of the sentences were used in more than one list. These sentences were recorded by a native male Kannada speaker using the Praat software (Boersma & Weenink, 2005).

An eight talker speech babble noise was used to generate sentences with different SNRs. In each list first sentence at was at +20 dB SNR and SNR was reduced in 5 dB steps for the subsequent sentences. Thus in each list, first sentence was at +20 dB SNR, second sentence was at +15 dB SNR, third sentence was at +10 dB SNR, fourth sentence was at +5 dB SNR, fifth sentence was at 0 dB SNR, sixth sentence was at -5 dB SNR and last sentence was at -10 dB SNR. These SNRs encompass the range of normal to severely impaired performance in noise. Sentences used were high probability items for which the key words are somewhat predictable from the context. Each sentence has five key words that are scored as correct/incorrect. These sentences were presented at 70 dB HL through a personal computer. The listener's task was to repeat the sentences presented and each correctly repeated keyword was awarded one point for a total possible score of 35 points per list. To calculate SNR at which 50% scores are obtained, the formula given below as mentioned in the study by Avinash, Methi and Kumar (2009) was used.

SNR at which 50% scores are obtained = 22.5 - (total words correct)

CHAPTER 4

RESULTS AND DISCUSSION

The present study aimed to compare the temporal resolution abilities and speech perception in noise in children born in families with musical background and children born in families without musical background. The temporal resolution abilities were measured using Temporal Modulation Transfer Function (TMTF) and Gap Detection Threshold (GDT) test. Temporal Modulation Transfer Function was tested for six different modulation frequencies (4 Hz, 8 Hz, 16 Hz, 32 Hz, 64 Hz & 128 Hz) for both the ears. Gap Detection threshold was also estimated for the individual ears separately. Quick Speech Perception in Noise – Kannada was obtained separately for both the ears. A total of 40 subjects participated in the present study, who were classified into 2 groups i.e., children born in families with musical background and the other group consisted of children born in families without musical background. These subjects were further classified into 4 subgroups with 10 subjects each.

- Children born in families without musical background and underwent no formal musical training- denoted as 'S_A'
- Children born in families without musical background but underwent formal musical training- denoted as 'S_B'
- Children born in families with musical background but underwent no formal musical training- denoted as 'S_1'
- Children born in families with musical background and underwent formal musical training- denoted as 'S_2'

The data were 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 different parameters of the three tests for both ears separately.
2. For temporal resolution tests such as TMTFs and GDT, multivariate analysis of variance (MANOVA) was administered to compare the parameters across all the four groups.
3. For the parameters which showed significant results under MANOVA, pair wise group comparison was done using Duncan's post hoc test.
4. Paired T-test was used to compare the ear effects among the groups.
5. For Quick Speech Perception in Noise – Kannada (QuickSPIN), Kruskal-Wallis test was administered to compare the parameters across all the four groups.
6. For the parameters which showed significant results under Kruskal-Wallis test, pair wise group comparison was done with the help of Mann-Whitney U test.
7. Wilcoxon Signed Rank test (for pair wise comparison) were done to compare the parameters within the group.

Temporal Resolution

Temporal Modulation Transfer Function:

Table 4.1

Mean and standard deviation (SD) of TMTF for the four groups at different modulation frequencies.

Modulation frequencies	Groups							
	S_A		S_B		S_1		S_2	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
4 Hz (R)	-22.85	1.89	-24.83	1.52	-22.84	1.67	-25.57	1.27
4 Hz (L)	-22.7	1.37	-25.05	1.35	-23.17	1.77	-25.80	1.21
8 Hz (R)	-19.97	1.98	-20.89	1.86	-20.11	2.00	-21.53	1.49
8 Hz (L)	-20.53	1.45	-21.17	1.41	-20.34	1.84	-21.17	1.50
16 Hz (R)	-14.90	1.34	-17.84	1.49	-15.02	1.33	-18.54	1.16
16 Hz (L)	-15.00	1.63	-17.97	1.20	-15.17	1.52	-18.70	1.39
32 Hz (R)	-12.20	1.68	-12.95	1.59	-12.56	1.51	-14.47	1.42
32 Hz (L)	-12.05	1.20	-13.11	1.60	-12.86	1.49	-14.57	1.37
64 Hz (R)	-6.65	1.65	-10.15	1.12	-7.05	1.57	-10.39	1.18
64 Hz (L)	-6.82	1.38	-10.28	1.05	-7.29	1.43	-10.67	1.27
128 Hz (R)	-3.38	0.86	-6.84	1.24	-3.99	0.85	-7.25	1.43
128 Hz (L)	-4.30	1.15	-6.69	1.37	-4.17	0.86	-7.32	1.37

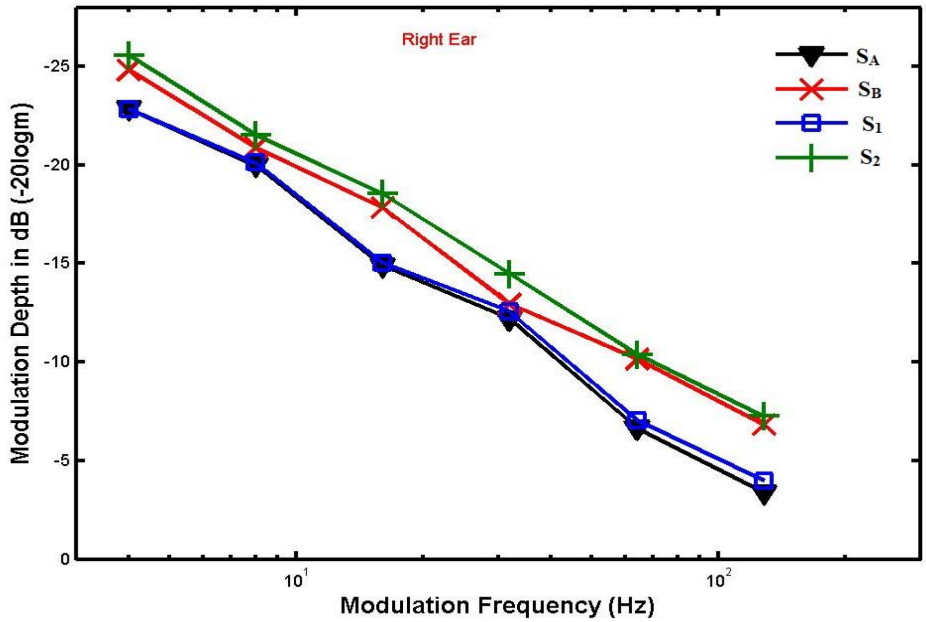

Figure 4.1. Mean values of the temporal modulation transfer function across different modulation frequencies of all the four groups for right ear.

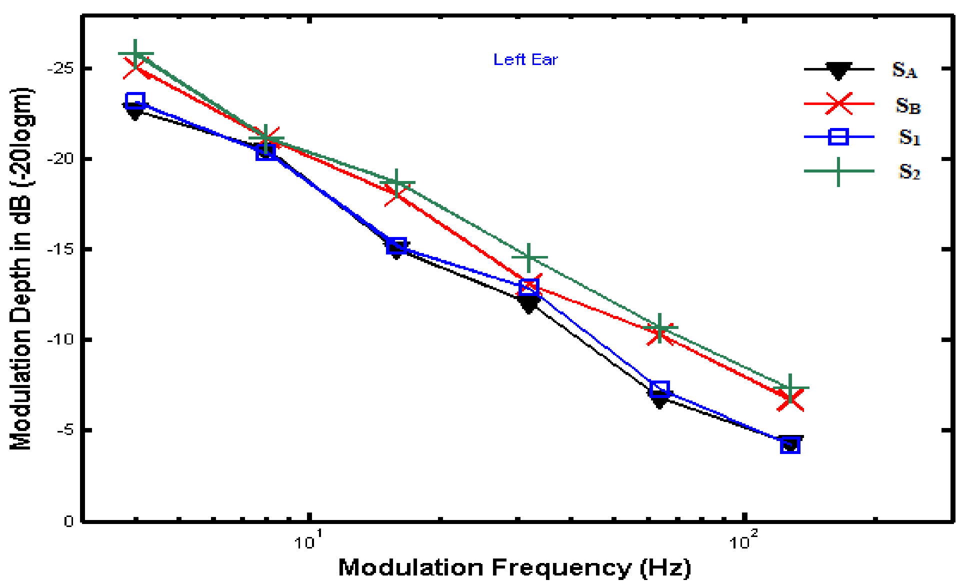

Figure 4.2. Mean values of the temporal modulation transfer function across different modulation frequencies of all the four groups for left ear.

Temporal modulation transfer function was derived from modulation detection thresholds for 6 different modulation frequencies (4 Hz, 8Hz, 16 Hz, 32 Hz, 64 Hz, and 128 Hz) for both the ears separately, for all the four groups. Table 4.1 shows the descriptive statistics (mean & SD) of the Modulation thresholds of all the six modulation frequencies across the four groups. The same are graphically depicted in figures 4.1 and 4.2 for right and left ears respectively. Different modulation frequencies (in Hz) were represented in the abscissa and the modulation detection thresholds or modulation depths (as 20 log m) were represented in the ordinate. It was observed that the group which had family background of music along with musical training (S_2) showed better temporal modulation detection thresholds in both ears, when compared with other groups. For all the groups the modulation detection threshold worsened as the modulation frequency increased.

Across Group Comparison:

Multivariate analysis of variance (MANOVA) was administered to compare the parameters across all the four groups. It revealed statistically significant difference for 4 Hz for both ears ($F(3,36) = 7.50$, $p<0.05$), 16 Hz for both ears ($F(3,36) = 19.84$, $p<0.05$), 32 Hz for both ears ($F(3,36) = 4.11$, $p<0.05$), 64 Hz for both ears ($F(3,36) = 19.99$, $p<0.05$) and 128 Hz for both ears ($F(3,36) = 26.26$, $p<0.05$) but statistically no significant difference was present for 8 Hz for both ears ($F(3,36) = 1.54$, $p>0.05$).

The results of MANOVA revealed that there is statistically significant difference between the scores among the groups. In order to find out which groups are statistically different, Duncan’s post hoc test was administered. At 4 Hz, 16 Hz, 32 Hz, 64 Hz and 128 Hz, for both ears, the groups S_1 and S_A showed statistically no difference at 5% level of significance. Similarly, the groups S_2 and S_B showed statistically no difference at 5% level of significance.

Within Group Comparison:

One-way repeated measure ANOVA was used for within group comparison. The results revealed that there was a statistically significant difference for all the frequencies of group S_A (F (5, 45 = 197.70, p<0.05), group S_B (F (5, 45) = 252.92, p<0.05), group S_1 (F (5, 45 = 210.41, p<0.05), and group S_2 (F (5, 45 = 292.35, p<0.05) for both ears.

Paired t- test was carried out to compare the ear effects within the groups at all the frequencies used for TMTFs. It revealed that there was no statistically significant difference between the ears at each frequency for all the groups at 5% level of significance.

Gap Detection Threshold (GDT) test:

Gap detection threshold (GDT) test was administered for both ears separately to find the minimum temporal gap, the subject could identify. GDT test was done for all the four groups. Mean and standard deviation (SD) of gap detection threshold for both the ears of the four sub groups are shown in Table 4.2.

Table 4.2

Mean and SD of Gap Detection Threshold (GDT) for both ears of all the groups.

	Groups							
	S_A		S_B		S_1		S_2	
	Mean(ms)	SD	Mean(ms)	SD	Mean(ms)	SD	Mean(ms)	SD
GDT (R)	3.68	0.30	3.37	0.36	3.35	0.25	3.11	0.32
GDT (L)	3.45	0.28	3.17	0.39	3.23	0.29	2.87	0.33

Descriptive statistical analyses showed that GDT was better for the group S_2 than compared to the other groups for both ears. Group S_2 was having a GDT of 3.11 ± 0.32 for right ear and 2.87 ± 0.33 for left ear whereas for group S_A , the GDT was 3.68 ± 0.30 for right ear and 3.45 ± 0.28 for left ear. For all the groups the GDT was better for left ear than compared to right ear as shown in figure 4.3.

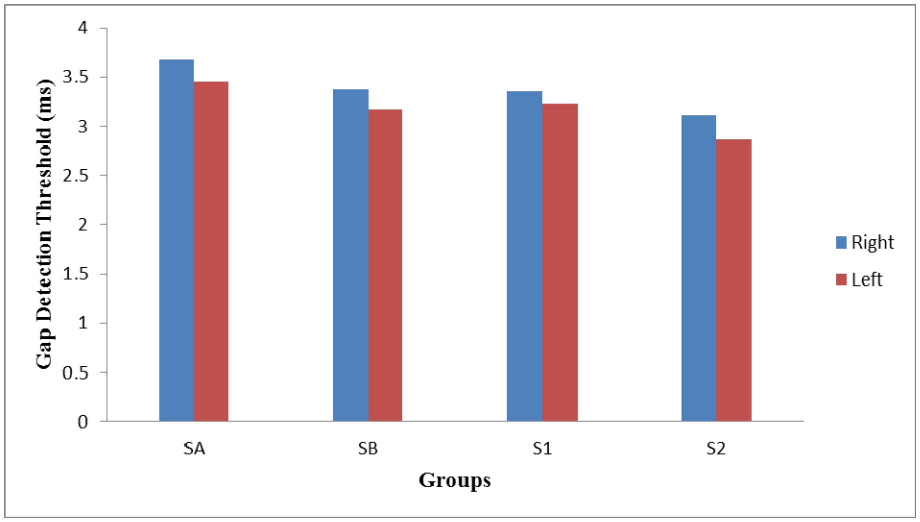

Figure 4.3. Mean values of Gap Detection Threshold for both ears across four groups.

Across Group Comparison:

Multivariate analysis of variance (MANOVA) was administered to compare the parameters across all the four groups. It showed statistically significant difference across the four groups ($F (3, 36) = 5.975$, $p<0.05$) for both the ears. Similarly, there was a statistically significant difference across the four groups when both right ear ($F (3, 36) = 5.569$, $p<0.05$) and left ear ($F (3, 36) = 5.296$, $p<0.05$) were compared.

In order to find out which groups are statistically different Duncan's post hoc test was administered. For right ear, the GDT for group S_A showed statistically significant difference when compared to the other groups at 5% level of significance. For left ear, the GDT for group S_2 showed statistically significant difference when compared to the other groups at 5% level of significance.

Within Group Comparison:

When all the groups are considered there was a statistically significant difference between the GDT thresholds for both the ears ($F (1, 36) = 36.45$, $p<0.05$). Paired t- test was carried out to compare ear effects within the groups. Each group was separately analyzed for ear differences. For group S_A, left ear GDT thresholds were better than right ear [$t (9) = 2.91$, $p<0.05$]. For group $S_{B,}$ left ear GDT thresholds were better than right ear [t (9) = 4.74, p<0.05]. For group S_1, left ear GDT thresholds were better than right ear [t(9) = 2.16, p<0.05)] and for group S_2, left ear GDT thresholds were better than right ear [t(9) = 3.08, p<0.05)].

Speech Perception in Noise

Quick Speech Perception in Noise – Kannada (QuickSIN):

Table 4.3

Mean and standard deviation (SD) of QuickSIN – Kannada of both the ears for all the groups.

	Groups							
	S_A		S_B		S_1		S_2	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Right ear	-2.20	2.62	-5.70	1.87	-1.90	1.64	-5.50	2.21
Left Ear	-3.10	2.31	-5.80	1.70	-2.00	1.71	-6.20	2.35

The descriptive statistics (Mean & SD) of the QuickSIN – Kannada for both ears across the four groups are shown in table 4.3. The mean values show the ability to perceive 50% of speech in presence of noise. It was found that the group S_2 perceived 50% of speech in presence of noise at most negative signal to noise ratios (SNR) when compared to other groups.

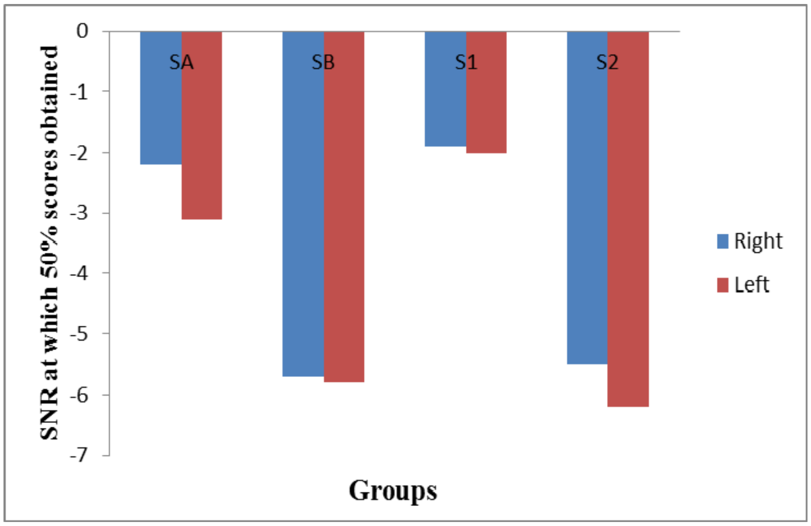

Figure 4.4 Mean values of SNR at which 50% scores are obtained for both the ears across four groups.

Across Group Comparison:

The descriptive statistics showed slightly higher standard deviation for all the groups. Hence, Krusikal Wallis test was administered to compare the SNR's across the 4 groups. For both right ear, $\chi^2_{(3)}$= 20.62, $p<0.05$ and left ear, $\chi^2_{(3)}$= 20.31, $p<0.05$, the results were statistically significant across the 4 groups.

Mann-Whitney U test was done to compare the SNR at which 50% scores are obtained between 2 groups. When groups S_A and S_B were compared, both right and left ears showed statistically significant difference, $|Z| = 3.14$ and $|Z| = 2.63$; $p<0.05$ respectively. When groups S_A and S_1 were compared, both right and left ears showed statistically no significant difference, $|Z| = 0.86$ and $|Z| = 1.47$; $p>0.05$ respectively. When groups S_A and S_2 were compared, both right and left ears showed statistically significant difference, $|Z| = 2.70$ and $|Z| = 2.72$; $p<0.05$ respectively.

When groups S_B and S_1 were compared, both right and left ears showed statistically significant difference, $|Z| = 3.50$ and $|Z| = 3.40$; $p<0.05$ respectively. When groups S_B and S_2 were compared, both right and left ears showed statistically no significant difference, $|Z| = 0.03$ and $|Z| = 0.23$; $p>0.05$ respectively. When groups S_1 and

S_2 were compared, both right and left ears showed statistically significant difference, $|Z| = 3.15$ and $|Z| = 3.44$; $p<0.05$ respectively.

Within Group Comparison:

Ear wise comparison was made within each group through Wilcoxon signed rank test. For group S_A, the SNR at which 50% scores obtained are statistically significant difference between right and left ears $|Z| = 2.12$; $p<0.05$. For group S_B, the SNR at which 50% score are obtained showed statistically no significant difference between right and left ears $|Z| = 0.27$; $p>0.05$. For group S_1, the SNR at which 50% score obtained showed statistically no significant difference between right and left ears $|Z| = 1.0$; $p>0.05$. For group S_2, the SNR at which 50% score obtained showed statistically no significant difference between right and left ears $|Z| = 1.84$; $p>0.05$.

Discussion

The results of temporal modulation transfer function across the four groups revealed statistically significant difference for the modulation frequencies like 4 Hz, 32 Hz, 64 Hz and 128 Hz, across different groups except 8 Hz. In the absence of the specific literature on TMTF in musicians it is difficult to explain the result found in the present study which has shown no difference at 8Hz rate. If tests to tap auditory working memory were included in the study those would have thrown light on providing explanations to results. Additionally it would have contributed to know the underlying processes involved. But in general, according to Ishll, et al. (2006), when random gap detection test was administered on musicians and non-musicians, the gap detection thresholds were better in trained musicians when compared to non-musicians. This helps to draw the conclusion that the temporal resolustion abilities are better in musicians when compared to non-musicians. The performance by the group S_2 and S_B were similar and statistically different from the groups S_1 and S_A. Therefore, the musical training along with the family background helps in shaping the temporal resolution abilities.

For GDT test, there was a statistically significant difference across the 4 groups. These results are in agreement with the study by Monteiro et al (2010), where it was concluded that musicians had better temporal resolution abilities when compared to non-musicians and the years of experience was a factor in deciding about the temporal resolution ability. Studies also reported that initiation of musical training in younger age also matters for the better abilities. At the same time, the influence of musical family background contributes to the better performance. According to Ohinshi et al (2001), music training can induce functional reorganization of the cerebral cortex. Therefore, musical training initiated early in life, before the age of seven contributes to the development of primary auditory cortex and more precisely the planum temporale. When the GDT was compared between the two ears within the group there was a statistical significant difference at 5 % level of significance. The performance in the left ear was better compared to that of the right ear. This could be an evidence to show that there is a better functional and anatomical organization of the right hemisphere in musicians.

For QuickSIN – Kannada, The group S_2 required the lowest SNR to perceive 50% of speech in the presence of noise. According to a study 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. When ear differences were compared, the group S_2 and S_B showed no ear differences while the groups S_1 and S_A showed better performance of left ear than right ear indicating that musical background and training stimulates and enhances the utility of both the hemispheres. Hence, the present study warrants exploring in depth and elucidating the reason for the better auditory performance.

CHAPTER 5

SUMMARY AND CONCLUSIONS

Many studies have reported that musicians have better auditory perceptual skills when compared to non-musicians. There are many studies in literature which have documented that musical training improves basic auditory perceptual skills resulting in enhanced behavioral (Jeon & Fricke 1997; Koelsch et al., 1999; Oxenham et al., 2003; Tervaniemi et al., 2005; Micheyl et al., 2006; Rammsayer & Altenmuller 2006) and neurophysiological responses (Brattico et al., 2001; Pantev et al., 2001; Schneider et al., 2002; Shahin et al., 2003, 2007; Tervaniemi et al., 2005; Kuriki et a.l, 2006; Kraus et al., 2009). Musicians' life long experience of detecting melodies from background harmonies can be considered as a process analogous to speech perception in noise. Studies report that musicians had a more robust sub- cortical representation of the acoustic stimulus in the presence of noise (Kraus et al., 2009). Musical practice not only enhances the processing of music related sounds but also influences processing of other domains such as language (Marques et al., 2007; Moreno et al., 2009; Parbery-Clark et al., 2009a; Schon et al., 2004, 2008). Because of their musical training, musicians have learned to pay more attention to the details of the acoustic stimulus than non-musicians. However, there is a dearth of literature to observe whether those advantages exhibited by them transcend the generations.

The present study aimed to find the temporal resolution abilities of children born in musical background and children born in non-musical background. Also, to find the speech perception abilities in the presence of background noise for the same group. A total of 40 subjects participated in the present study. An informal interview was administered to all participants, in order to get the information regarding their training in musical field and also familial musical background. The participants were classified into 4 subgroups based on their family background and musical training. Each group consisted of 10 subjects. Temporal resolution abilities were found out using Temporal Modulation Transfer Function (TMTF) and Gap Detection Threshold (GDT) test. Quick

Speech Perception in Noise – Kannada (QuickSIN) was administered to check speech perception in the presence of noise. All these tests were administered at 40 dB SL or at most comfortable level, for both ears separately.

The results from the present study showed that the temporal resolution abilities and the ability to perceive speech in the presence of noise were better in children with musical training than compared to children without musical training. Musical training as a factor has contributed to better performance whether or otherwise of the family background. Secondly, the family background with musical training has shaped the auditory processing skills to the finest level. In the context of no family background, musical training has yielded good performance. Finally in the absence of family background and musical training the auditory processing skills have not shown significant differences.

Implications of the Study:

- Music training can be used as a potential remediation strategy for children requiring language training and auditory processing disorders.
- Can be implemented in Hearing Aid technology for musicians with hearing loss to improve their speech perception.
- Genetic predisposing factors for musical skills can be explored.
- Music training can be advocated for all children irrespective of whether they have musical background or otherwise to facilitate speech perception in adverse listening conditions and probably to provide the advantage of enhanced auditory working memory.
- Further research on the influence of auditory working memory on speech perception in noise is warranted.
- To add information to the literature.

Acknowledgements

We would like to thank the Director, AIISH, for providing the funds for this project. We are also grateful to staff and participants of Vijaya Vittala Vidyashala, Mysore.

REFERENCES

Alain, C., & Brenstein, L .J. (2008). From sound to meaning: The role of attention during auditory scene analysis. *Current Opinions of Otolaryngologica: Head and Neck Surgery, 16*, 485-489.

American National Standards Institute (1991). *Maximum Ambient Noise Levels for Audiometric Test Rooms*. (ANSI S3. 1-1991). New York: American National Standards Institute.

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.

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

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

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.

Boersma, P. & Weenick, D. (2005). *Praat doing phonetics by computer* (version 4.6.09) from www.praat.org

Bradlow, A. R., Kraus, N., Hayes, E. (2003). Speaking clearly for children with learning disabilities: sentence perception in noise. *Journal of Speech Language and Hearing Research, 46*, 80-97.

Brandler, S., & Rammsayer, T.H. (2003). Differences in mental abilities between musicians and non-musicians. *Psychology of Music, 31 (2)*, 123-138.

Brandt, J., Rosen, J.J. (1980). Auditory phonemic perception in dyslexia: categorical identification and discrimination of stop consonants. *Brain and language, 9*, 324-337.

Brattico, E., Naatanen, R., & Tervaniemi, M. (2001). Context effects on pitch perception in musicians and nonmusicians: Evidence for event related potentials. *Music Perception, 19*, 199-222.

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

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

Carhart, R., & Jerger, J. F. (1959). Preferred method for clinical determination of pure tone thresholds, *Journal of Speech and Hearing Disorders, 36*, 476-483.

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

Chandrasekaran, B., Krishnan, A., & Gandour, J. T. (2009). Relative influence of musical and linguistic experience on early cortical processing of pitch contours. *Brain and Language, 108*, 1-9.

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 (5)*, 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*, 35-66.

Foregeard, M., Schlaug, G., Norton, A., Rosam, C., Iyengar, U., & Winner, E. (2008). The relation between music and phonological processing in normal-reading children and children with dyslexia. *Music Perception, 25*, 383-390.

Fujioka, T., Trainor, L & Ross, B. (2008). Simultaneous pitches are encoded separately in auditory cortex: An MMN study. *Neuroreport, 19*, 361-368.

Fujioka, T., Trainor, L. J., Ross, B., et al. (2005). Automatic encoding of polyphonic melodies in musicians and nonmusicians. *Journal of Cognitive Neuroscience, 17*, 1578-1592.

Fujioka, T., Trainor, L., Ross, B., Kakigi, R., & Pantev, C. (2004). Musical training enhances automatic encoding of melodic contour and interval structure. *Journal of Cognitive neuroscience, 16(6)*, 1010-1021.

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

Gordon-Salant, S., Fitzgibbons, P. J. (1995). Recognition of multiply degraded speech by young and elderly listeners. *Journal of Speech Language and Hearing Research, 38*, 1150-1156.

Ho, Y.C., Cheung, H. C., & Chan, A. S. (2003). Music training improves verbal but not visual memory: Cross sectional and longitudinal explorations in children. *Neuropsychology, 17*, 439-450.

Ishll, C., Arashiro, A. M., & Pereira, L. D. (2006). Ordering and temporal resolution professional singers and amateurs tuned and un-tuned. *Pro-Fono, 18(3)*, 285-292.

Jackobsen, L. S., Cuddy, L. L., & Kilgour, A. R. (2003). Time tagging: A key to musicians' superior memory. *Music Perception, 20*, 307-313.

Jakobson, L. S., Lewycky, S.T., Kilgour, A. R., & Stoesz, B. M. (2008). Memory for verbal and visual material in highly trained musicians. *Music Perception, 26*, 41-55.

Jeon, J. Y., & Fricke, F. R. (1997). Duration of perceived and performed sounds. *Psychology of Music, 25*, 70–83.

Koelsch, S., & Mulder, J., (2002). Electric brain responses to inappropriate harmonies during listening to expressive music. *Clinical Neurophysiology, 113*, 862-869.

Koelsch, S., Schroger, E., & Tervaniemi, M. (1999). Superior attentive and pre-attentive auditory processing in musicians. *Neuroreport, 10*, 1309– 1313.

Koelsch, S., Schulze, K., Sammler, D., et al. (2008). Functional architecture of verbal and tonal working memory: An fMRI study. *Human Brain Mapping, 30*, 859-873.

Koffka, K. (1935). Principles of Gestalt Psychology. New York: Harcourt, Brace and World.

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

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.

Kuriki, S., Kanda, S., & Hirata, Y. (2006). Effects of musical experience on different components of MEG responses elicited by sequential piano-tones and chords. *Journal of Neuroscience, 26*, 4046–4053.

Lee, K. M., Skoe, E., Kraus, N., & Ashley, R. (2009). Selective subcortical enhancement of musical intervals in musicians. *Journal of Neurosciences, 29*, 5832-5840.

Levit, H. (1971). Transformed Up-Down methods in psychoacoustics, *Journal of Acoustical Society of America, 49*, 467-477.

Magne, C., Schon, D., & Besson, M. (2006). Musician children detect pitch violations in music and language better than non-musician children: Behavioural and electro-physiological approaches. *Journal of Cognitive Neurosciences, 18*, 199-211.

Marques, C., Moreno, S., & Casstro, S. L. (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: Behavioural and electrophysiological evidence*.* *Journal of Cognition and Neuroscience, 19*, 1453–1463.

Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psycho acoustical training on pitch discrimination. *Hearing Research, 219*, 36–47.

Monteiro, R. A., Nascimento, F. M., Soares, C. D., Ferreira, M. D. (2010). Temporal resolution abilities in Musicians and no musicians. *International Archives of Otolaryngology, 14(3),* 302-308.

Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S.L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year old children: more evidence for brain plasticity. *Cerebral Cortex 19*, 712-723.

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)*, 75-760.

Oxenham, A. J., Fligor, B. J., Mason, C. R., & Kidd, G. (2003). Informational masking and musical training. *Journal of Acoustical Society of America, 114*, 1543–1549.

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L.E., Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature, 392*, 811–814.

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.

Patel, A. D. (2003). Language, music, syntax and the brain. *Natural Neuroscience, 6*, 674-681.

Patel, A. D. (2007). *Music, Language and the Brain*. Oxford: Oxford University Press.

Rajashekar, B & Vyasamurthy, M. N. (1976). *Development and standardization of a picture SRT test for adults and children in Kannada*. Unpublished Master's Dissertation, University of Mysore.

Rammsayer, T., & Altenmuller, E. (2006). Temporal information processing in musicians and non-musicians. *Music Perception, 24*, 37-48.

Scheffers, M., (1983). *Sifting Vowels: Auditory Pitch Analysis and Sound Segregation*. Unpublished Doctoral dissertation. Groningen. The Netherlands: Groningen University.

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

Schlaug, G., Norton, A., Overy, K., & Winner, E. (2005). Effects of music training on the child's brain and cognitive development. *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.

Schon, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology, 41*, 341–349.

Shahin, A. J., Roberts, L. E., & Pantev, C. (2007). Enhanced anterior-temporal processing for complex tones in musicians. *Clinical Neurophysiology, 118*, 209–220.

Shahin, A., Bosnyak, D.J., Trainor, L.J., Robertsm, L.E. (2003). Enhancement of neuro-plastic P2 and N1c auditory evoked potentials in musicians. *Journal of Neuroscience, 23*, 5545–5552.

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.

Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., Schorger, E. (2005). Pitch discrimination accuracy in musicians vs. non-musicians: an event-related potential and behavioural study. *Experimental Brain Research, 161*, 1- 10.

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding speech prosody: Do music lessons help? *Emotion, 4*, 46-64.

Treisman, A. M. (1964). Verbal cues, language and meaning in selective attention. *American Journal of Psychology, 77*, 206-219.

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.

Ziegler, J. C., Pech-Georgel, C., George, F., Alario, F. X., Lorenzi, C. (2005). Deficits in speech perception predict language learning impairment. *Proceedings of National Academic Science USA, 102*, 14110-14115.

Temporal Modulation Transfer Function (TMTF)

Two stimuli, unmodulated white noise and sinusoidally amplitude modulated (SAM) white noise of 500 ms duration, with a ramp of 20 ms were used. The stimuli was generated using a 32 bit digital to analog converter with a sampling frequency of 44.1 KHz and were low pass filtered with a cut-off frequency of 220 Hz. The Modulated signal was derived by multiplying the white noise by a dc-shifted sine wave. The depth of modulation was controlled by varying the amplitude of modulating sine wave. Modulation depth was varied between 0 to -30 dB (where 0 dB is equal to 100% modulation depth and -30 dB is equal to 0% modulation). Six different modulation frequencies were used (4 Hz, 8 Hz, 16 Hz, 32 Hz, 64 Hz, & 128 Hz).

The stimuli were presented at 40 dB SL (with reference to PTA) or at comfortable level. The stimuli were presented to the participants through headphones. The subjects' task was to discriminate between modulated and the unmodulated noise till they were able to identify the difference.

Three interval alternate forced choice methods (3IAFC) were used. On each trial, un-modulated and modulated stimuli were successively presented with an inter-stimulus interval of 500 ms. Modulation depth was converted into decibels [20 log 10(m), where 'm' refers to the depth of modulation]. A step size of 4 dB was used initially and then reduced to 2 dB after two reversals. This procedure provides an estimate of the value of amplitude modulation necessary for 70.7 % estimate of correct responses (Levitt, 1971). The mean of eight reversals in a block of 14 will be taken as threshold.

Gap Detection Threshold (GDT)

Gap detection test consisted of a standard stimulus of 750 ms duration Gaussian noise with a silence of standard duration placed at its temporal center. The variable stimulus had variable gap duration and the length of its gap is changed as a function of the subject's performance. All noises had a 0.5 ms cosine ramp at both onset and offset. Three Interval Alternate Forced Choice Method (3IAFC) methods were used to obtain the gap detection threshold. It consisted of three blocks of white noise, one of which contained gaps of variable duration. The subjects' task was to identify the gap and to detect which block of noise was having the gap in it. The presentation level of the stimulus was 40 dB SL (with reference to PTA) or most comfortable level, monaurally. Each time when the subject detected the gap embedded in noise correctly, the size of the gap was reduced and test was continued till the subject could trace the smallest gap. The minimum gap that the subject detected was considered as the gap detection threshold. The gap detection thresholds were obtained for both the groups.