BRAINSTEM CORRELATES OF SPEECH PERCEPTION IN NOISE: CARNATIC MUSICIANS VS. NON-MUSICIANS

Jubin Vinod
Register No. 10AUD036

A Dissertation Submitted in Part Fulfillment of Final Year
Master of Science (Audiology),
University of Mysore, Mysore.

All India Institute of Speech and Hearing
Manasagangothri, Mysore-570006
May 2012.

CERTIFICATE

This is to certify that this dissertation entitled **“Brainstem Correlates Of Speech Perception In Noise: Carnatic Musicians Vs. Non-Musicians”** is the bonafide work submitted in part fulfillment for the Degree of Master of Science (Audiology) of the student with Registration No.: 10AUD036. 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 the award of any other Diploma or Degree.

Dr. S. R. Savithri

Director

Mysore **All India Institute of Speech and Hearing,**

May, 2012 **Manasagangothri, Mysore- 570006.**

CERTIFICATE

This is to certify that this dissertation entitled **“Brainstem Correlates Of Speech Perception In Noise: Carnatic Musicians Vs. Non-Musicians”** has been prepared under my supervision and guidance. It is also certified that this has not been submitted earlier in other University for the award of any Diploma or Degree.

Dr. K. Rajalakshmi

(Guide)

Reader in Audiology,

Department of Audiology,

All India Institute of Speech and Hearing,

Manasagangothri, Mysore- 570006

Mysore

May, 2012

DECLARATION

This is to certify that this Master's dissertation entitled **"Brainstem Correlates Of Speech Perception In Noise: Carnatic Musicians Vs. Non-Musicians"** is the result of my own study under the guidance of Mr. K. Rajalakshmi, Lecturer in Audiology, Department of Audiology, All India Institute of Speech and Hearing, Mysore, and has not been submitted earlier in other University for the award of any Diploma or Degree.

Mysore Register No. 10AUD036.

May, 2012

ACKNOWLEDGEMENT

Every man is a product of two elements: nature defines him as who he is by birth and nurture defines who he becomes in life. It has been a long way for me since March 31st 1989 to the present. The journey has been defined by moments of pleasure and pain, joy and grief, victory and defeat. Needless to say, it would have been impossible for me to get this far had it not been for some significant people in my life who have always been a constant source of support and guidance for me.

I would like to begin by my mother, who ushered me into this world and taught me to remain true to myself, no matter what the circumstances and that the qualities of love, kindness and understanding have no equal in this world.

I would like to thank my father for being my moral compass in life, for teaching me the value of honesty and integrity and to always stick to my guns, no matter how bad things get.

I express my sincere gratitude to my guide Dr. Rajalakshmi for being a source of continuous support and guidance throughout my dissertation and my life at AIISH. Madam. You are more than just a lecturer to me. You are a source of enlightenment and joy to all who surround you. Your humility and kindness, despite your position *would put even the greatest of intellectuals to shame. You are truly a role model without equal.*

I would like to thank Dr. S. R. Savithri, Director, All India Institute of Speech and Hearing, Mysore, for permitting me to carry out this study.

I would like to thank the HOD, Department of Audiology, Dr. Animesh Barman., for permitting me to use the test equipments to conduct the study.

My sincere gratitude to my lecturers at AIISH- Asha Madam, Rajalakshmi mam, Sujeet Sir and Sandeep Sir, for imparting valuable knowledge to me and laying the foundations of my professional life. Thank you for being my teachers.

I would like to specially thank Ganapathy Sir, Hemanth Sir, Jijo Sir and Anoop B.J. for their valuable insights which helped shape my dissertation.

My thanks to Arun Raj Sir, Anthony Sir, Sharath sir, Prashanth Sir, Jithin Sir and for helping me out in my data collection.

I would like to thank all participants who had participated in this study, who had taken time off their busy schedules to help me out.

I would like to thank Ms. Vasathalakshmi and Mr. Santosh for their help in statistical analysis and for their invaluable suggestions.

During mu life in AIISH, I had a number of great seniors to look up to- my heartfelt thanks to Jijo P.M., Kuppuraj 'The Hunk' S, Mohammed Ismail, Giridhar 'G-Man' Krishan, Ratnakar 'Rising Ratna' Y., Vivek 'Moloch' Mandal, Rohit H., Anoop 'Samaajam' O.T., Nirmal Sugathan, Wishly Mammen, Nike G., Srikar V, Mariechi, Ranjini Chechi, Nisha Chechi, Merin Chechi, Shafna Chechi, Sneha Chechi and Rexy Chechi for making life here fun.

My class has (and always will be) number one for me. My heartfelt thanks to Sneha (my oldest and closest friend), Saravanan (The King), Akshay (The Karate Kid), Spoorthi, Sahana, Mythri, Apoorva, Deepashriee (Pops), Arpitha (Nagara), Deepika, Arya, Divya, Jasmine, Chandan (The Katy), Prajeesh, Hijas, Jonathan (my constant source of entertainment), Rohit (Singham), Reuben (Chatur) and all my classmates for the wonderful times and the unforgettable memories. Thank you for being with me through all these years….

My heartfelt thanks to all my lovely juniors- Akbar, Mohandas, Ujjwal (my dinner partner), Sindhu, Indira, Ansu, Sandhya, Himanshu (Salman), Thareeque, Jim, Abhishek and others who complete the circle of my life at AIISH.

This list of acksnowledgements would not be complete if I failed to mention one person who, through the years has come to mean a lot to me. It has been a tumultuous roller-coaster ride over the past few years, though at the end of the day you will always be the angel who so brightened my life with your presence. Words cannot even begin to express how much you mean to me and how much I am going to miss you. However. I do know, deep in my heart that even though fate may choose to separate us now, we will be together again, sometime, someplace, forever and always.

TABLE OF CONTENTS

Chapter	Title	Page No.
1	Introduction	1-7
2	Review of Literature	8-25
3	Method	26-34
4	Results and Discussion	35-71
5	Summary and Conclusion	72-73
	References	74-84

LIST OF TABLES

Table No.	Title	Page No.
Table 3.1	Recording Protocol For Speech ABR in Quiet and in Noise	31
Table 4.1	Results of Mixed ANOVA to check for interactions of Latency measures across Groups and Conditions	39
Table 4.2	Mean and Standard Deviation for Peak Latencies of Non-Musicians and Musicians in Quiet	40
Table 4.3	Mean and Standard Deviation for Peak Latencies of Non-Musicians and Musicians in Noise	41
Table 4.4	Paired t-test results for Latencies of waves of Non-Musicians in Quiet and in Noise	42
Table 4.5	Paired t-test results for Latencies of waves of Musicians in Quiet and in Noise	42
Table 4.6	One Way MANOVA for Latencies of Waves in Quiet	43
Table 4.7	One Way MANOVA for Latencies of Waves in Noise	44
Table 4.8	Results of Mixed ANOVA to check for interactions of Latency measures across Groups and Conditions	47
Table 4.9	Mean and Standard Deviation for Peak Amplitudes of Non-Musicians and Musicians in Quiet	48
Table 4.10	Mean and Standard Deviation for Peak Amplitudes of Non-Musicians and Musicians in Noise	49

Table 4.11	Paired t-test results for Amplitudes of waves of Non-Musicians in Quiet and in Noise	50
Table 4.12	Paired t-test results for Amplitudes of waves of Musicians in Quiet and in Noise	51
Table 4.13	One Way MANOVA for Amplitudes of Waves in Quiet	52
Table 4.14	One Way MANOVA for Amplitudes of Waves in Noise	53
Table 4.15	Results of Mixed ANOVA to check for interactions of Formant Amplitude measures across Groups and Conditions.	56
Table 4.16	Mean and Standard Deviation for Formant Amplitudes of Non-Musicians and Musicians in Quiet	57
Table 4.17	Mean and Standard Deviation for Formant Amplitudes of Non-Musicians and Musicians in Noise	58
Table 4.18	Paired t-test results for Formant Amplitudes of waves of Non-Musicians in Quiet and in Noise	59
Table 4.19	Paired t-test results for Formant Amplitudes of waves of Musicians in Quiet and in Noise	60
Table 4.20	One Way MANOVA for Formant Amplitudes of Waves in Quiet	61
Table 4.21	One Way MANOVA for Formant Amplitudes of Waves in Noise	61
Table 4.22	Results of Independent Sample t-Test for comparing SPIN Scores across the 2 groups	64

Table 4.23	Mean and Standard Deviation of SPIN Score In Musicians and Non-Musicians	64
Table 4.24	Results of Pearson's Correlation: Correlation of SPIN Scores with Latency, Amplitude and Formant Amplitude in Quiet for Non-Musicians	65
Table 4.25	Results of Pearson's Correlation: Correlation of SPIN Scores with Latency, Amplitude and Formant Amplitude in Quiet for Musicians	67
Table 4.26	Results of Pearson's Correlation: Correlation of SPIN Scores with Latency, Amplitude and Formant Amplitude in Noise for Non-Musicians	69
Table 4.27	Results of Pearson's Correlation: Correlation of SPIN Scores with Latency, Amplitude and Formant Amplitude in Noise for Musicians	70

LIST OF FIGURES

Figure No.	**Title**	**Page No.**
Figure 3.1	Spectral and temporal aspects of the Speech stimulus /da/	29
Figure 3.2	FFT representing the energies at the fundamental frequency and its harmonics	34
Figure 4.1	Speech-Evoked ABR in response to 40 msec /da/ in a Non-Musician in quiet and in noise	36
Fig.4.1	Speech-Evoked ABR in response to 40 msec /da/ in a Musician in quiet and in noise	37

CHAPTER 1

INTRODUCTION

Music can be defined as an art form that arranges sounds in a fashion that follows certain natural principles and provides that special inner feeling of happiness and contentment. It is important to note that the basic principles are natural and thus the theory of music is only an attempt by man to rationally explain what is already beautiful. As a fringe benefit, this rationalization helps in understanding the inherent beauty of music and creates increasingly higher levels of appreciation in the listener. Music perception is a complex, cognitively demanding task that taps into a variety of complex brain functions underlying acoustic analysis, auditory memory, auditory scene analysis and processing of musical syntax.

Through years of sensory-motor training, often beginning in early childhood, musicians develop an expertise in their instrument of specialization or mastery over their voice. In the course of training, musicians increasingly learn to attend to the fine-grained acoustics of musical sounds. These include pitch, timing and timbre, the three basic components into which any sound that reaches the human ear — including music or speech — can be broken down. Pitch refers to the organization of sound on an ordered scale ranging from low to high pitch and is a subjective percept of the frequency of the sound. Timing refers to specific landmarks in the sound (for example, the onset and offset of the sound) and timbre refers to the quality of the sound — a multidimensional attribute that results from the spectral and temporal features in the acoustic signal. Attention to these components is emphasized during music training. For example, a violinist is trained to pay particular attention to pitch cues to effectively tune the violin, an instrumentalist playing in an orchestra has to have a keen sense of timing cues and a conductor needs to rely on timbre cues to differentiate the contribution of various instruments. A variety of studies have found that musical training improves auditory-perceptual skills resulting in enhanced behavioural (Jeon and Fricke, 1997; Koelsch, Schroger, & Tervaniemi, 1999; *Micheyl Delhommeau*, *Perrot* and Oxenham.,2006; Rammsayer and Altenmuller, 2006; Tervaniemi, Kruck, Baene, Schröger, Alter & Friederici, 2009) and neurophysiological (Brattico, Naatanen and Tervaniemi, 2001; Pantev et al., 2001; Schneider, Scherg, Dosch, Specht, Gutschalk & Rupp, 2002; Shahin , Bosnyak , Trainor, Roberts & Larrey, 2003; Tervaniemi, Kruck, Baene, Schröger, Alter & Friederici, 2005; Kuriki, Kanda, & Hirata, 2006; Kraus., Skoe, Parbery-Clark & Ashley, 2009) responses. These findings are attributable to their intensive training, during which musicians learn to pay more attention to details of acoustic stimuli than non-musicians (Musacchia, Sams, Skoe, & Kraus, 2007).

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. Thus it is only reasonable to assume that the benefits that musicians have in processing music would also extend to speech stimuli. A number of research studies have shown 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).

Musical practice has also been found to result in enhancement of other verbal and non-verbal skills such as auditory attention (Strait, Kraus, Parbery-Clark, & Ashley, 2010), auditory stream segregation (Beauvois & Meddis, 1997), processing emotion in speech (Strait, Kraus, Skoe,& Ashley, 2009a), working memory (Chan, Ho, & Cheung, 1998; Forgeard, Winner, Norton, & Schlaug, 2008) and processing of prosody and linguistic features in speech (Bidelman, Gandour, & Krishnan, in press; Chandrasekaran, Krishnan, & Gandour, 2009; Wong, Skoe, Russo, Dees & Kraus, 2007).

Of special note is the enhanced ability of musicians to extracting relevant signals from a complex soundscape (e.g., the sound of their own instrument in an orchestra). Speech perception in noise is a complex task that requires the segregation of target signals from a competing background noise. To complicate matters, the noise also degrades the signal particularly by disrupting the perception of rapid spectro-temporal changes (Brandt and Rosen, 1980). Poor performance in the task of speech perception in noise is seen in individuals with hearing impairment (Gordon- Salant and Fitzgibbons, 2005) and language-based learning disabilities (Bradlow, Kraus, & Hayes, 2003; Ziegler, Pech-Georgel, George, & Lorenzi, 2005) whereas musicians demonstrate better performance than non-musicians (Parbery-Clark, Skoe and Kraus, 2009). It was hypothesized that a musician’s long-term experience with musical stream segregation would transfer to the homologous task of speech perception in noise. Parbery-Clark, Skoe, Lam and Kraus (2009) found a distinct speech in noise advantage for musicians, as measured by two standardized tests of hearing in noise (HINT, Hearing in-noise test;

QuickSIN). Musicians showed superior working memory and performed better on a frequency discrimination task. 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 ability, leading to an overall advantage in speech perception in noise in musicians.

All these enhanced abilities in musicians may be related to structural and functional enhancements seen at different levels of their nervous system. For instance, highly trained musicians exhibit unique anatomical, functional and event-related specializations as opposed to non-musicians. For instance, musicians have more neural cell bodies (grey matter volume) in the auditory, motor and visuo-spatial areas of the brain (Gaser and Schlaug, 2003) and also have more axonal projections that connect the right and left hemispheres (Schlaug, Jancke, Huang, Staiger & Steinmetz, 1995). Professional instrumentalists, compared to amateurs or untrained controls, display greater activation in auditory areas such as Heschl's gyrus (Schneider, Scherg, Dosch, Specht, Gutschalk & Rupp, 2002) and the planum temporale (Ohishi, Matsuda, Asada, Aruga, Hirakata & Nishikawa, 2001) in response to sound. Some event related potentials from auditory cortical areas such as the N1, N1c and P2 were found to be more robust in musicians as compared to non-musicians (Pantev et al., 1998; Shahin et al., 2003). All these anatomical enhancements are seen to translate into improved auditory and cognitive skills as is evidenced by various studies. The intensive practice over the years has been attributed to bring about neuroplastic changes in the practitioner as is evidenced in many research studies (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). One of the mechanisms used to explain the findings of music-induced experience dependent plasticity at the level of the brainstem is increased efficiency of top-down predictive coding (Strait, Kraus, Parbery-Clark, & Ashley, 2010).

The auditory system is composed of a number of neural structures that are interconnected via bottom-up (ascending) as well as top-down (descending) pathways. Perceiving speech in noisy environments is a complex task involving higher-level cognitive and lower-level sensory processing (Nahum, Nelken, & Ahissar,2008). The signal has to be delivered to higher cortical structures with enough fidelity that it can be decoded as being meaningful (Hickok & Poeppel, 2007). For this to happen, the impact of background noise needs to be minimized. Recent studies have suggested an important role for the feedback (top-down) pathways in fine-tuning the auditory signal at early stages of auditory processing (Luo, Wang, Kashani, & Yan, 2008). Such top-down influences back-project all the way to the cochlea through the medial olivocochlear bundle (MOCB). These authors have said that feedback initiated by the higher (cortical) structures is transferred to the lower (brainstem) structures via the efferent auditory system. This results in an enhanced selectivity of sound features at the lowest levels of the auditory system which is important for higher-level structures to distinguish relevant information in the signal from irrelevant details.

The human auditory brainstem response (ABR) has been used as an index of brainstem encoding of speech stimuli (Chandrasekaran & Kraus, 2010; Skoe & Kraus, 2010). The ABR to a plosive speech syllable (for e.g., /da/) consists of an onset response that marks the consonant burst, and a frequency-following response (FFR) that reflects phase-locked responses to the consonant-vowel transition as well as the vowel portion of the stimulus. Stop consonants are particularly vulnerable to the deleterious effects of noise due to their transient nature (Brandt & Rosen, 1980). Since the FFR preserves spectral information up to about ~2000 Hz and reflects neural timing in the order of milliseconds, it can therefore be used to examine the fidelity of the brainstem representation of spectral and timing information. It has been found that the addition of background noise delays the timing of brainstem responses (Cunningham, Nicol, Zecker, & Kraus, 2000; Cunningham, Nicol, Zecker, Bradlow, & Kraus, 2001;) and reduces spectral magnitude.

There is evidence from studies using speech-evoked ABR that music training modulates the effect of background noise on subcortical auditory representation (Parbery- Clark, Skoe,& Kraus, 2009). Musicians show less degraded brainstem representation of speech relative to non-musicians, as evidenced by faster neural timing, enhanced spectral representation, and better stimulus-to-response correlations. Though the differences between musicians and non-musicians are present even in quiet backgrounds (Musacchia et al., 2007), it is in the presence of background noise that the differences in spectral representation between musicians and non-musicians are large, suggesting that musical experience protects against the debilitating effects of background noise (Parbery-Clark, Skoe, et al., 2009). Thus timing and spectral features are preserved at the level of the brainstem to a greater extent due to musical experience and these enhancements translate into a better performance on the task of speech perception in noise. The speech-evoked ABR is hence considered to be a reliable indicator of the biological basis of speech perception in noise.

NEED FOR THE STUDY

Parbery-Clark, Skoe & Kraus (2009) have shown that despite the well documented disruptive effects of noise, western musicians have a more robust sub-cortical encoding of speech in the presence of background noise. This enhanced neural representation of speech has been seen to result in an enhanced perception of speech in the presence of background noise (Parbery- Clark, Skoe & Kraus, 2009; Parbery- Clark, Skoe, Lam & Kraus, 2009).

There is a need to execute a similar study in Indian musicians since no such studies have been done to verify whether the same enhancements exist in Carnatic musicians.

AIM OF THE STUDY

The aims of the current study were to:

- To verify whether trained Carnatic musicians show better perception of speech in the presence of background noise as compared to non-musicians
- To verify the presence of enhanced subcortical encoding of speech stimuli in trained Carnatic musicians relative to non-musicians

CHAPTER 2

REVIEW

The research done on musicians has revealed 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 aspects like linguistic skills, working memory, temporal abilities, perception of emotions and the ability to perceive speech in the presence of noise. These enhancements are directly related to the presence of various structural and functional differences between musicians and their untrained peers.

Changes in structural & functional aspects of nervous system

Highly trained musicians exhibit unique anatomical, functional and event-related specializations as opposed to non-musicians. For instance, musicians have more neural cell bodies (grey matter volume) in the auditory, motor and visuo-spatial areas of the brain (Gaser and Schlaug, 2003) and also have more axonal projections that connect the right and left hemispheres (Schlaug et al., 1995). Perhaps as a consequence of these enhancements, professional instrumentalists, compared to amateurs or untrained controls, display greater activation in auditory areas such as Heschl's gyrus (Schneider, Scherg, Dosch, Specht, Gutschalk & Rupp, 2002) and the planum temporale (Ohishi, Matsuda, Asada, Aruga, Hirakata & Nishikawa, 2001) in response to sound. Musical training also promotes plasticity in somatosensory regions; with string players demonstrating larger areas of finger representation than untrained controls (Elbert, Pantev, Wienbruch, Rockstroh & Taub, 1995).

The Heschl's gyrus, which contains the primary auditory area, was found to be larger in musicians than non-musicians and its size correlated with musical proficiency as measured by psychometric evaluation (Schneider et al., 2002). It has also been observed that the left planum temporale, which is important for the processing of complex sounds, is relatively larger than the right planum temporale in professional musicians (Schlaug, G., 2001). With respect to the integrity of directionally 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, Nagy, Skare, Forsman, Forssberg & Ullén, 2005).

Functionally, the brain responses of adult musicians and non-musicians also differ as measured by EEG and MEG. Some event related potentials from auditory cortical areas such as the N1, N1c and P2 were found to be more robust in musicians as compared to non-musicians (Pantev, Roberts, Schulz, Engelien, Almut, Ross & Bernhard., 1998; Shahin , Bosnyak , Trainor, Roberts & Larrey, 2003).

Recent studies have suggested that playing a musical instrument "tunes" the neural activity of structures peripheral to the cortex (Musacchia, Sams, Skoe & Kraus, 2007; Wong, Skoe, Russo, Dees, & Kraus, 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 enhancements corresponded to stimulus features that may be particularly important for processing music. One such enhancement is observed with the frequency following response (FFR), which is thought to be generated primarily in the inferior colliculus and consists of phase-locked neural responses whose inter-spike intervals occurs at the fundamental frequency (F0) of a sound (Hoormann, Falkenstein, Hohnsbein, & Blanke, 1992; Krishnan, Xu, Gandour & Cariani, 2005). Since F0 is understood to underlie the percept of pitch, this response is hypothesized correspond to the ability to accurately encode acoustic cues for pitch. Enhanced encoding of this aspect of the stimulus would clearly be beneficial during perception of musical pitch. Accordingly, the previous studies have shown the presence of larger peak amplitudes at F0 as obtained via Fast Fourier Transform analysis of the FFR and better pitch tracking in musicians relative to non-musicians. Another enhancement was observed in the wave delta (which occurs ~8ms post-acoustic onset) of the brainstem response to sound onset, which has been hypothesized to be important for encoding stimulus onset (Musacchia et al., 2007). Stimulus onset is an attribute of music important for the denotation of instrument attack and rhythm, and correspondingly the authors observed earlier wave delta responses in musicians than non-musicians. Additionally, this enhancement of FFR and wave delta in musicians was observed in response to both music and speech stimuli and was the most prominent when the 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 similar stimulus characteristics.

Musacchia, Strait & Kraus (2008) studied the relationship between evoked potentials and musical experience. They simultaneously recorded 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 an auditory stimulus in isolation, when the subjects watched a video of a male speaker saying /da/ in conjunction with the auditory stimulus and when they viewed the video in isolation. The analysis was focused on the comparison of measures of the speech evoked brainstem response that were previously reported as being enhanced in musicians with well established measurements of cortical activity (e.g., P1-N1-P2 complex). It was found that recent musical training improves auditory memory and shapes P1-N1 response and encoding of F0. The correlation between electrophysiological and behavioral measures suggests that performance on complex auditory task is related to the strength of the P1-N1 response. The musicians performed better on the behavioral tests and showed 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 in response to pitch, timing, and timbre features of music, relative to non-musicians (Pantev et al., 2001).

Musicians had a statistically stronger correlation between brain-related and behavioral 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 to explain experience-induced plasticity, which is known to be expressed in AEPs in adults (Tremblay, Kraus, McGee, Ponton & Otis, 2001). Shahin, Roberts & Trainor, (2004) compared AEPs evoked by pure tones, violin and piano tones in young 4- to 5- year old children with musical experience and age matched non-musician controls. 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 were measured in independent cohorts of children without musical training between 4 to 15 years of age in response to the same stimuli, which provided a baseline against which the data from the experimental group was compared. The 4 to 5 year old children with musical experience were found to have larger amplitudes of P1, N1, and P2 responses than their non-musician peers. 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. The comparison of piano-evoked N1 and P2 responses in 4- to 5-year-old musicians (most of whom were pianists) to cross sectional findings suggest that musical experience may have advanced the developmental trajectory for sounds pertaining to the instrument of training.

In sequential stimuli, a wrong note occurring occasionally 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, Trainor , Ross, & Kakigi**,** 2004). In terms of polyphonic music, altered notes in either of simultaneous melodies elicit MMN responses that were found to be larger in musicians than non-musicians (Fujioka, Trainor, Ross, Kakigi & Pantev, 2005).

Thus, all these anatomical enhancements are seen to translate into improved auditory and cognitive skills as is evidenced by various studies that have examined the behavioural adaptations of musicians as an effect of training. Among these enhanced skills, of particular note are the enhanced perception of pitch and the improvement seen in working memory.

Enhancements in Auditory and Cognitive 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 involves the use of pitch contours and intervals to communicate melodies and tone centers. Pitch patterns in speech convey prosodic information which is used by listeners to identify indexical information (information about the speaker’s intention as well as emotion and other social factors). In tonal languages, changes in pitch are also used lexically (in differentiating between words) (e.g., In Mandarin: /ma/ (high level) means ‘mother’, /ma/ (high rising) means ‘hemp’ , /ma/ (low falling rising) means ‘horse’, ma (high falling) means ‘scold’) .

Musicians have been known to display enhanced processing of prosodic and linguistic pitch. Musicians show an enhanced ability to detect subtle incongruity in prosodic pitch and also show consistent neural differences relative to non-musicians (Besson, Schon, Moreno, Santos & Magne, 2007; Magne, Schon, & Besson, 2006). Differences between musicians and non-musicians are evident even during pre-attentive stages of auditory processing (Musacchia et al., 2007; Wong & Perrachione, 2007). Frequency following responses (FFR), which are neural responses originating from 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 variations provide meaningful information) were seen to be better at processing pitch contours (even in a non-linguistic context) than native speakers of English (Bent, Bradlow, & Wright, 2006). Physiologically, Mandarin speakers show more robust encoding of the pitch content of Mandarin sounds at cortical and sub-cortical levels of their auditory system, suggesting that language experience fundamentally changes the neural circuitry of the auditory pathway (Krishnan, Xu, Gandour, & Cariani, 2005). Musicians showed superior cortical representation of linguistic pitch in a non-native language relative to non-musicians (Krishnan et al., 2005).In these individuals, 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 representation of non-native pitch in the lower levels of the auditory system. Using synthetic speech stimuli that contained F0 contours representative of citation forms of Mandarin and Thai lexical tones, it was seen that experience-dependent brainstem mechanisms for pitch representation, as reflected in pitch-tracking accuracy and pitch strength, are more sensitive in tonal (Chinese, Thai) than in non-tonal (English) language speakers (Krishnan and Gandour, 2009). Findings suggest that, for a non-native language musicians showed superior cortical representation of linguistic pitch 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, these 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). The enhancements in musicians are 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, Lam and Kraus., 2009). As a whole, these studies demonstrate that musicians show a distinct advantage in the early auditory processing of speech features.

A number of studies have evidenced enhancements in musicians 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 elements. 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, Norton and Schlaug, 2008; Jakobson, Lewycky, Kilgour, & Stoesz, 2008; Parbery-Clark, Skoe, Lam, and Kraus, 2009), attention (Strait, Kraus, Parbery-Clark, & Ashley, 2010; Tervaniemi, Kruck, Baene, Schröger, Alter, & Friederici, 2009), and executive function (Bialystok & DePape, 2009) abilities. Musicians are also significantly better than non-musicians in auditory stream segregation, presumably due to their music training (Beauvois & Meddis, 1997; Zendel & Alain, 2009). Chan, Ho & Cheung (1998) 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.

To summarize, relative to non-musicians, musicians have shown enhanced verbal memory, improved sensory representation of speech features including pitch, timing, and timbre, enhanced stream segregation, working memory, attention, and executive skills. All these skills underlie successful perception of speech in noisy backgrounds.

Enhanced perception of Signals 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 to the skills that subserve successful perception of speech in noise. Perception of speech in noise is a complex task requiring perceptual cue detection, stream segregation and working memory. A distinct advantage is seen in musicians on the task of perception of speech in noise as measured by standardized tests of hearing in noise such as Hearing-In-Noise test & QuickSIN (Parbery-Clark, Skoe, Lam and Kraus., 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 ability. These correlations strongly suggest that such intensive practice fine-tunes both cognitive and sensory abilities, leading to an overall advantage in speech perception in noise in musicians. Musicians were also found to perform 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 than on the spatial segregation of the sound sources.

Perception of speech in noise may also be affected by changes in central auditory processing. Degenerative changes due to aging were found to affect the ability to process pitch cues (Helfer & Vargo 2009). The ability to perceive speech in the presence of the noise in increasingly poorer SNRs (0 dB, -5 dB & -10 dB) was seen to increase as a function of the experience of the musicians. 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, unpublished dissertation 2011).

In order to find the effect of musical experience on the neural representation of speech signals in noise, Parbery-Clark, Skoe & Kraus (2009) compared sub-cortical neurophysiological responses to speech in quiet and in 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 robust responses in background noise than control group. Musicians also had earlier response onset timing and better phase locking to the temporal waveform and stimulus harmonics than non-musicians. They also found that earlier response onset 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 sub-cortical representation of speech in the presence of background noise, which may contribute to musician's behavioral advantage for speech in noise perception. Musicians also exhibited more faithful encoding the steady state portion of a 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. A 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 sub-cortical sensory representations of sounds through the enhancement of the target signal and the suppression of irrelevant competing background noise via the efferent system (Suga, Zhang and Yan, 1997; Zhang, Suga and Yan, 1997; Luo, Wang, Kashani & Yan, 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 sub-cortical signal quality, resulting in a more robust representation of the target acoustic signal in noise. The sub-cortical encoding of the $F0$ of the speaker is an important factor in the perception of speech in noise. The $F0$, along with 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 partly dependent on the *F*0, as demonstrated by the enhanced discrimination of vowels with greater *F*0 separation between concurrent words (Assmann & Summerfield, 1987) and sentences (Brokx, Nooteboom and Cohen, 1982).

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 were found to have enhanced cortical responses to their primary instrument suggesting that their specific listening and training experiences modulate their neural responses to specific timbres (Pantev et al., 2001; Margulis, Mlsna, Uppunda, Parrish & Wong, 2009).

Role of Neural Plasticity

All these findings can be attributed to the enhanced neural plasticity at the level of thee brainstem and the cortex.

Evoked potentials were used to analyze the development of the auditory brainstem response to clicks and speech sounds in children between the ages of 3 and 12 years. The neural response to a click stimulus showed similar response latency across all age groups, which was in agreement with previously established reports (Salamy, 1984; Gorga, M.P., Kaminski, Beauchaine, Jesteadt, & Neely, 1989; Ponton, Eggermont, Coupland, & Winkelaar, 1992; Abdala and Folsom, 1995). In contrast, peak latency measurements throughout the brainstem response to speech were significantly longer for 3 to 4 year-old children compared with 5 to 12 year-olds. Systematic age related changes were also seen in the latency of speech evoked Binaural Interaction Component.(BIC). The latency of BIC obtained using speech stimuli in children aged between 6.11-7.11 yrs were significantly prolonged than in children aged between 8-12 years. However, there was no such difference in latency of BIC for clicks. These findings indicate that the BIC continues to develop till about 8 years of age.(Sonitha, unpublished dissertation 2011).

Evidence for the experience-dependent nature of neural plasticity in humans is derived from literature on statistical learning, which describes the manner in which the auditory system reacts to frequently occurring sounds. At the level of the inferior colliculus, the neural populations rapidly adjust their firing patterns based on the statistical distribution of the sounds encountered, and these adjustments engender improved coding accuracy for sounds occurring the most commonly (Dean, Harper and McAlpine, 2005), even in an on-line manner. Krishnan et al. (2005) found that native Mandarin speakers had increased accuracy in pitch tracking compared to native English-speaking adults. Studies carried out by Musacchia et al. (2007) and Wong, et al., (2007) found enhanced brainstem encoding of the F0 in musicians. These studies speak of the effect of long-term auditory experiences initiated in childhood on the neural encoding of sounds. Moreover, short-term training has been shown to improve brainstem timing in children with learning problems (Russo et al., 2005),

In a study conducted by Shahin, Bosnyak, Trainor, Roberts (2003), highly skilled violinists, pianists and non-musician controls passively attended to violin tones, piano tones, and pure tones matched in fundamental frequency to the musical tones. Compared with non-musician controls, both musician groups evidenced larger N1c (latency of appoximately 138 msec) and P2 (latency of appoximately 185 msec) responses to the three types of stimuli. As is seen in training studies with non-musicians, N1c enhancement was expressed preferentially in the right hemisphere, where auditory neurons may be specialized for processing of spectral pitch.

Thus, the changes seen in musicians as a function of experience may be attributed to enhanced neural plasticity.

Neural Bases of Speech Perception in Noise

The auditory system is composed of a number of neural structures that are interconnected via bottom-up (ascending) as well as top-down (descending) pathways. Perceiving speech in noisy environments is a complex task involving higher-level cognitive and lower-level sensory processing (Nahum, Nelken, & Ahissar, 2008). The signal has to be delivered to higher cortical structures with enough fidelity that it can be decoded as being meaningful (Hickok & Poeppel, 2007). For this to happen, the impact of background noise needs to be minimized.

Recent studies have suggested an important role for the feedback (top-down) pathways in fine-tuning the auditory signal at early stages of auditory processing (Luo, Wang, Kashani, & Yan, 2008). These authors address three important principles underlying automatic sound selection by top-down feedback pathways. Specifically: a) feedback is initiated by higher level structures (i.e., cortex), b) efferent pathways carry this information to lower-level structures such as the auditory brainstem, and c) selectivity arises at the earliest stages of processing. This selectivity is important for higher-level structures to distinguish relevant information in the signal from irrelevant details.

Sub-Cortical Processing of Speech in Noise

The human auditory brainstem response (ABR) has been used as an index of brainstem encoding of speech stimuli (Chandrasekaran & Kraus, 2010; Skoe & Kraus, 2010). The ABR to a plosive speech syllable (for e.g., /da/) consists of an onset response that marks the consonant burst, and a frequency-following response that reflects phase-locked responses to the consonant-vowel transition as well as the vowel portion of the stimulus. The ABR to the consonant-vowel stop syllable has been extensively studied in typical and clinical populations (Tzounopoulos & Kraus, 2009). Stop consonants are particularly vulnerable to the deleterious effects of noise due to their transient nature (Brandt & Rosen, 1980). Because the FFR preserves spectral information up to about ~2000 Hz and reflects neural timing in the order of milliseconds, it can therefore be used to examine the fidelity of the brainstem representation of spectral and timing information. In general, the addition of background noise delays the timing of brainstem responses (Cunningham, Nicol, Zecker, & Kraus, 2000; Cunningham, Nicol, Zecker, Bradlow, & Kraus, 2001; Hall, 1992) and reduces spectral magnitude.

Recent studies have examined experience-dependent plasticity in the representation of speech in background noise. Music training modulates the effect of background noise on subcortical auditory representation (Parbery- Clark, Skoe,& Kraus, 2009). Musicians show less degraded brainstem representation of speech relative to nonmusicians, as evidenced by faster neural timing, enhanced spectral representation, and better stimulus-to-response correlations. The differences between musicians and nonmusicians are present, albeit to a lesser degree even in quiet backgrounds (Musacchia et al., 2007). In background noise however, the differences in spectral representation between musicians and nonmusicians are large, suggesting that musical experience protects against the debilitating effects of background noise (Parbery- Clark, Skoe, and Kraus, 2009). Thus timing and spectral features are preserved to a greater extent due to musical experience.

Brainstem representation of speech in noise also has been examined in children with behavioural deficits in noise-exclusion. Relative to children who show good perception of speech in noise, those with noise-exclusion deficits show delayed brainstem response timing and poorer representation of pitch in background noise (Anderson, Skoe, Chandrasekaran, & Kraus, in press; Anderson, Skoe, Chandrasekaran, Zecker, & Kraus, 2010). Interestingly, these children do not differ in quiet conditions, revealing a biological basis for the behavioral deficits in noise-exclusion. Behavioral performance on hearing in noise tests is also associated with the brainstem differentiation of stop-consonants (ba/da/ga) (Hornickel, Skoe, Nicol, Zecker, & Kraus, 2009). Children who showed clear neural differentiation of the three contrastive stimuli at the level of the brainstem also demonstrated better speech in noise perceptual skills. Taken together, noise appears to blur the representation of timing and spectral elements important for speech perception in noise at the level of the brainstem.

Top-Down Influence on Early Auditory Processing

Higher-level auditory structures influence processing in lower-level structures via the efferent auditory network called the corticofugal pathways. Such top-down influences back-project all the way to the cochlea through the medial olivocochlear bundle (MOCB). The functioning of MOCB can be noninvasively examined in humans by measuring the suppression of evoked otoacoustic emissions, which are sounds generated within the cochlea in response to acoustic stimulation. Electrical stimulation of the auditory cortex can modulate MOCB activity in human participants (Perrot et al., 2006). To study the role of top-down modulation on speech in noise perception, MOCB activity was examined in young participants who underwent a training program to discriminate speech presented in noisy environments (De Boer & Thorton, 2008). Interestingly, an increase in MOCB activity correlated with speech in noise performance in good perceivers. In fact, learning outcomes could be predicted by MOCB activity. The authors conclude that corticofugal feedback plays an important role during listening in noisy environments. In the context of previous animal and human studies that have examined the corticofugal pathway, it is possible that top-down modulation improves signal quality at the auditory periphery by selectively amplifying relevant features of the signal, and inhibiting irrelevant features in the presence of background noise. Recent studies have argued that children with learning problems show a deficit in the ability to modulate early sensory encoding of speech features. In contrast, studies also have suggested that musicians show a superior ability to modulate sensory representation based on topdown cues.

Enhanced responses to native as well as non-native speech stimuli have been argued to be a result of an increased efficiency of the corticofugal network (Musacchia et al., 2007; Parbery-Clark, Skoe, et al., 2009; Wong et al., 2007).Musicians showed enhanced induced gamma-band activity (GBA), which is oscillatory brain activity in the 25 Hz-100 Hz range. Induced GBA is argued to reflect integration of top-down and bottom-up sensory processing (Trainor, Shahin, & Roberts, 2009). One year of music training in children has been shown to increase induced GBA relative to untrained participants (Shahin, Roberts, Chau, Trainor, & Miller, 2008). These authors argue that GBA changes reflect increased efficiency of top-down processes, and that music has dramatic effects on cognitive-sensory interaction.

Thus, as a combination of all these factors, musicians are seen to have an enhanced representation of stimulus features at the level of the brainstem, which is highly resistant to degradation in noisy environments and which translates into an enhanced performance on tasks of speech perception in noise.

CHAPTER-3

METHOD

The study was conducted with the aim of investigating the effect of musical training on the encoding of speech stimuli at the level of the brainstem by examining and comparing speech evoked ABR responses of both Carnatic musicians and non-musicians in quiet and in the presence of noise, and by relating these responses to the performance of the subjects in the task of speech perception in noise.

Participants

Participants of the present study were divided into two groups.

Group I: 15 native Kannada speakers (30 ears) with over 10 years of formal training in Carnatic music with normal hearing sensitivity in the age range of 20 to 50 years participated in the study.

Group II: 15 native Kannada speakers (30 ears) without any prior training in music with normal hearing sensitivity in the age range of 20 to 50 years participated in the study.

Participant selection criteria:

1. All the participants had normal hearing thresholds as evidenced by air conduction thresholds of less than or equal to 15 dB HL in the octave frequency range of 250 Hz to 8000 Hz and bone conduction thresholds of less than or equal to 15 dB HL in the octave frequency range of 250 Hz to 4000 Hz.
2. All the participants had Speech Identification Scores of > 90% at 40 dB(SL) with reference to Speech Recognition Thresholds in both ears.

3. All the participants had normal middle ear functioning as evidenced by tympanometry and reflexometry results.
4. Participants did not have any history of otological or neurological problems.

Instrumentation

Following equipments were used for the study:

1) Pure Tone Audiometer

A two channel OB922 audiometer with TDH-39 head phone coupled to impedance matched TDH 39 earphones with MX-41/ AR ear cushions and a bone vibrator (Radio ear B-71) was used to obtain pure tone threshold at different frequencies for both air conduction and bone conduction, as well as the speech recognition thresholds in quiet and in the presence of speech noise.

2) Immittance meter

A calibrated automatic Immittance meter with a visual display (Grason - Stadler GSI-TS) was used to rule out middle ear abnormalities. Each ear of the participant was tested for the type of tympanogram and presence or absence of acoustic reflexes.

3) Evoked potential system

An evoked potential system (Biologic Navigator Pro EP) was used to record both speech evoked ABR in quiet and in the presence of white noise.

4) Laptop

A Sony VPCEH25EN Laptop was connected to the audiometer to present the phonemically balanced wordlist in Kannada (Yathiraj and Vijayalakshmi, 2005) in the presence of ipsilateral speech noise.

Test environment

All the audiological evaluation and recording were carried out in a sound treated room. The ambient noise was within the permissible limits as recommended by ANSI (S3.1; 1991).

Test Stimulus for speech ABR:

The test stimulus which was used for speech evoked ABR in the present study was a synthesized /da/ syllable. The stimulus is available in evoked potential system with the BioMARK protocol. The /da/ stimulus is a 40 ms synthesized speech syllable produced using KLATT synthesizer (Klatt, 1980). This stimulus simultaneously contains the broad spectral and fast temporal information characteristic of stop consonants, and spectrally rich formant transitions between the consonant and the steady-state vowel. Although the steady-state portion is not present, the stimulus is still perceived as being a consonant-vowel syllable. The fundamental frequency (F0) 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 shows both the time and spectral domain of the stimulus used in the present study.

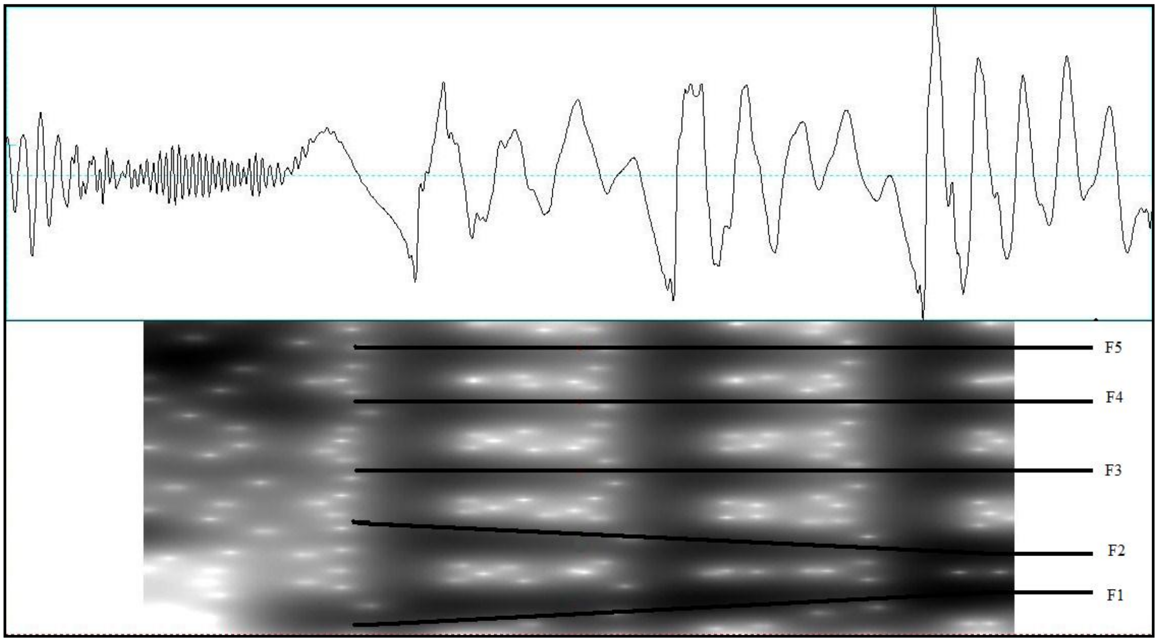


Figure 3.1 Spectral and temporal aspects of the Speech stimulus /da/ used in the present study. The top one represents the temporal details of the waveform whereas the bottom one depicts the spectral details.

The same /da/ syllable was simultaneously presented along with ipsilateral white noise for recording of speech ABR in the presence of noise. This feature was also available in the BioMARK protocol. The stimulus was presented at an SNR of 0dB.

Procedure

1)Pure tone audiometry

Behavioral air conduction and bone conduction thresholds were tracked using modified Hughson and Westlake procedure (Carhart & Jerger, 1959). Air conduction thresholds were obtained from 250Hz to 8 KHz and bone conduction thresholds were obtained from 250Hz to 4 KHz. Participants who had thresholds within 15 dB HL has further undergone immittance assessment in both the ears.

2) Tympanometry

Tympanometry was done to rule out pathology of middle ear using 226Hz probe tone. Immittance test was carried out by sweeping the pressure from +200 to -400 dapa. In reflexometry both ipsilateral and contralateral acoustic reflexes thresholds were measured for 500 Hz, 1000 Hz, 2000 Hz, and 4000Hz pure tone at the peak pressure.

3) Electrophysiological recording

Disc electrodes were placed on the recording sites with conduction paste and secured with skin tape. It was made sure that the impedance of each electrode was within <5 k Ω and inter electrode impedance was within <2 k Ω. The impedance of each electrode was also checked during testing, to make sure that patient movement did not cause any variation in the impedance. Participants were instructed to sit comfortably on a reclining chair and relax during the testing. They were instructed to close their eyes during the testing to avoid any artifacts. ABR were recorded twice for the reproducibility for both speech and non-speech stimuli.

Speech ABR was recorded using the protocol given in Table 3.1.

Table 3.1

Recording Protocol For Speech ABR in Quiet and in Noise

Stimulus, duration **Masker**	CV syllable /da/, 40 ms In Quiet- None In Noise- Ipsilateral White Noise (0 dB SNR)
Level	80 dB SPL
Filter band	70 to 2000 Hz
Rate	7.1/s
No of sweeps	2000
Transducer	BioLogic Insert ear phone
Polarity	Condensing + Rarefaction (Weighted Added)
Time window	64 msec which included a prestimulus time of 10 msec
Electrode montage	Non-inverting electrode: Vertex Inverting electrode: Test ear Mastoid Ground electrode: Low Forehead

Analysis of ABR recordings:

Analysis of Speech evoked ABR:

The electrophysiological brainstem responses to speech sound are a complex waveform, comprised of onset peaks as well as sustained elements. The following parameters were measured for each speech ABR recording:

a. *Latency and Amplitude of Onset and Transition peaks:*

The latencies of the two onset peaks, V and A, which mark the onset burst of the speech syllable were measured as was the amplitude of V peak.

The latencies of the transition peaks D, E, and F, which arise in response to the fundamental periodicity of the stimulus, were measured. For the FFR to be considered to be present, the fluctuation in the activity of the brainstem should repeat itself with a time period of approximately 10msec. The time period would correspond to the F0 (100Hz) of the stimulus frequency (Frequency= 1/ Time period). The three major peaks which repeated itself at the time period of 10msec were considered as D, E, and F. The amplitude of wave D, E and F were analysed.

c. *Pitch:*

The sustained FFR portion which occurs immediately after the onset response was subjected to Fast Fourier Transform (FFT) to obtain information regarding the spectral characteristics of the FFR (frequency and amplitude of spectral peaks). The average spectral amplitude was calculated for a frequency range from 103–120 Hz which encompasses the fundamental frequency (F0). FFT was performed on all speech evoked potentials using a custom made program run in MATLAB. The peak amplitude corresponding to F0 was also calculated using a custom made program file in the MATLAB platform. The frequency analysis was done from 11.4 to 40.6 msec. The sustained portion of the response (FFR) was passed through 103-120Hz band pass fourth order Butterworth filters in order to obtain the energy at F0. The Fourier analysis was then performed on the filtered signal. A subject's responses were required to be above the noise floor in order to be included in the analysis. This was performed by comparing the spectral magnitude of pre stimulus period to that of the response. If the quotient of the magnitude of F0 frequency component of FFR divided by the pre stimulus period was >1, the response was deemed to be above the noise floor.

d. *Harmonics:*

The Harmonics measure is a composite of the average spectral energy from two frequency bands: first formant (F1) 220 to 720 Hz, and high frequency (HF) 721–1154 Hz. F1 includes the harmonics of the stimulus that make up the most prominent frequencies of the first formant range in the analysis time of 11.4 to 40.6 msec. The HF range is composed of harmonics between the first and second formants (F1 and F2, respectively). The sustained portion of the response (FFR) was passed through 200 to 720 band pass fourth order Butterworth filters in order to obtain the energy at F1. Because higher formants are above the phase locking limits of the brainstem, no higher frequency ranges were included.

Figure-3.2 shows maximum amplitude in the F0 region i.e. around 103 to 125Hz. There is also some amount of energy in the F1 region i.e. from 220 to 720 Hz.

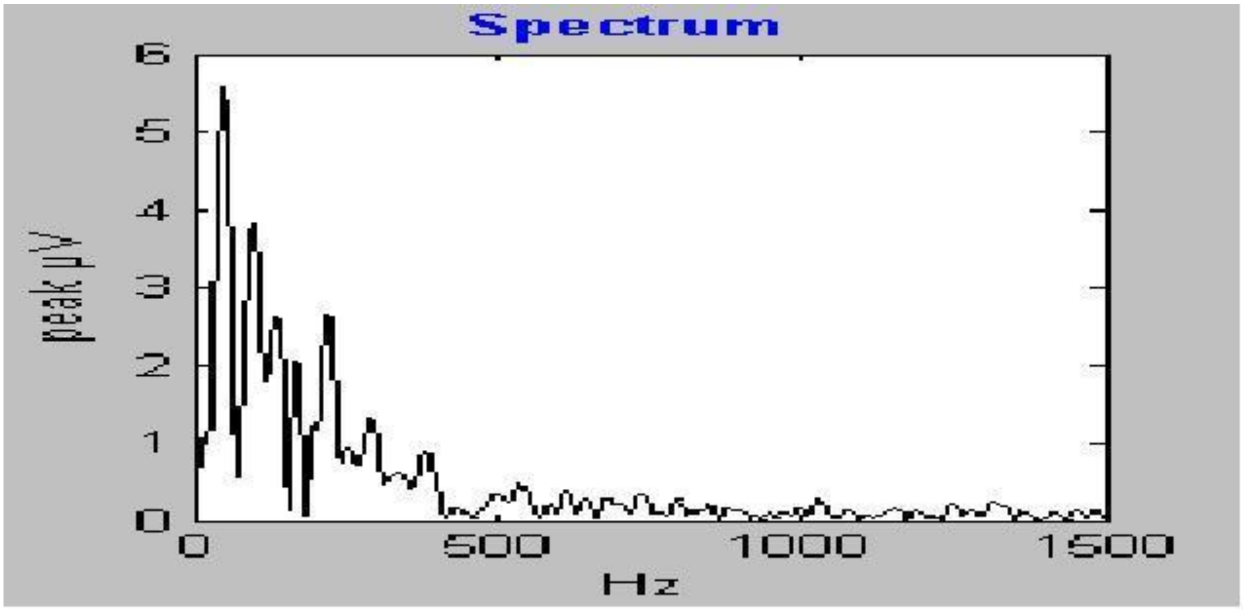


Figure 3.2 FFT representing the energies at the fundamental frequency and its harmonics.

CHAPTER 4

RESULTS AND DISCUSSSION

The present study was designed to examine the effect of musical training on the encoding of speech stimuli at the level of the brainstem. The speech evoked ABR responses of both Carnatic musicians and non-musicians were acquired in the absence and in the presence of noise (0 dBSNR). The responses were analyzed and were related to the performance of the subjects on the task of speech perception in noise.

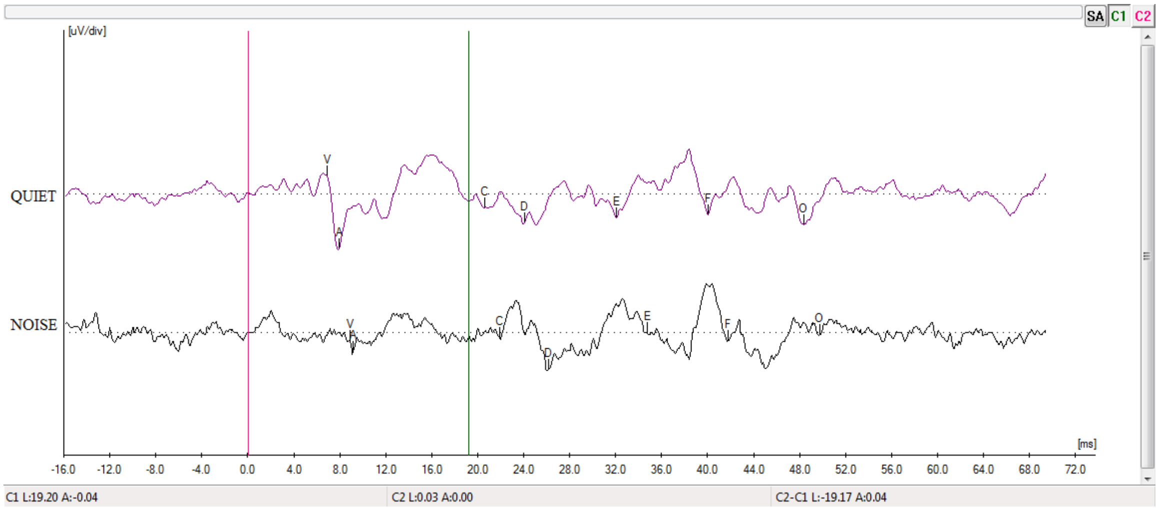
Fifteen individuals who had undergone formal musical training for a minimum of ten years and fifteen individuals who had no formal musical training constituted the experimental and control groups respectively. Speech-Evoked ABR was recorded from each individual in both conditions and these responses were subject to Fast Fourier Transform using MATLAB. The latencies and amplitudes of the waves V, D, E and F and the amplitudes of the encoded formants F0 (103-121 Hz), F1 (454-719 Hz) and F2 (721-11555 Hz) were subjected to statistical analysis.

The following statistical analyses were done to compare the data from the control and experimental groups:

- Descriptive statistics was done to obtain the mean and standard deviation for the SPIN scores, latencies and amplitudes of waves V, D, E and F as well as for the amplitudes of F0, F1 and F2 in the control and experimental groups obtained for both conditions.
- Mixed Analysis Of Variance (ANOVA) was carried out to check for the presence of main effects and interaction effects as a function of conditions, groups and variables (latency and amplitudes of peaks and formant amplitude)

- One-way Multiple Analysis of Variance (MANOVA) was carried out to check for significant differences between the groups for the different latency, amplitude and formant amplitude measures obtained in quiet and in noise.
- Paired t-test was carried out to compare the latency, amplitude and formant amplitude measures across quiet and noise in musicians and non-musicians.
- Independent sample t-test was carried out to verify whether the SPIN scores varied significantly across musicians and non-musicians
- Pearson’s correlation coefficient was calculated to measure the correlation between SPIN scores and the latency, amplitude and formant amplitude measures in quiet and in noise for both groups.

Speech ABR in Non- Musicians



C1 L:19.20 A:-0.04

C2 L:0.03 A:0.00

C2-C1 L:-19.17 A:0.04

Fig.4.1 Speech-Evoked ABR in response to 40 msec /da/ acquired in a Non-Musician in quiet and in noise (0dB SNR)

In quiet, the onset peaks V and the transition peaks D, E, and F were clearly visible in the speech evoked ABR of the non-musicians.

The morphology of the waves was noticeably poorer in noise, with peaks having reduced amplitude and delayed latencies. As in Fig 4.1, the V-A complex is almost eliminated in noise, though the transition waves are less affected. Similar findings were reported in studies by Russo, Nicol, Zecker, Hayes and Kraus (2004) & Russo, Nicol, Musacchia and Kraus (2004).

Speech ABR in Musicians

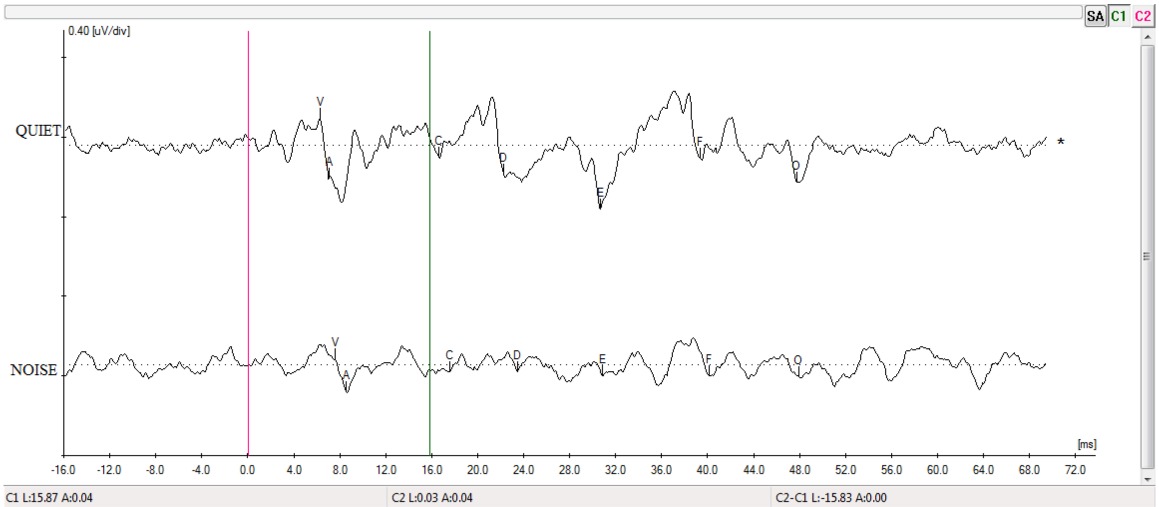


Fig.4.1 Speech-Evoked ABR in response to 40 msec /da/ acquired in a Musician in quiet and in noise (0dB SNR)

In quiet, the morphology of the Speech-Evoked ABR did not vary much from that seen in non-musicians. Though the waveform morphology was poorer in noise than in quiet, the waves were by and large better defined than in the corresponding waveforms of non-musicians. The V-A complex in particular is more clearly seen (Fig 4.2). This is in line with the findings of Parbery-Clark, Skoe and Kraus (2009).

Comparison of peak latencies

The latencies of the onset peak V and the transition peaks D, E and F were considered for analysis. The latencies of the peaks are related to the timing of the features of the stimulus (the onset and transition portions). The addition of background noise had been documented to result in delays in latencies of the peaks of ABR, indicating a disruption in timing of brainstem activity (Don and Eggermont, 1978; Cunningham et al., 2001; Russo et al., 2004). It has been hypothesized that the disruptive effects of noise on the representation of stimulus features may be limited by long-term musical training which can bring about enhancements of stimulus features at the sub cortical level via top down influences (Dean, Harper and McAlpine, 2005) mediated through the efferent auditory system (Luo, Wang, Kashani, & Yan, 2008).

The comparison of latency measures obtained in different conditions across the groups using mixed ANOVA (TABLE 4.1) reveals the presence of main effects of conditions and groups as well as interaction effects between groups, conditions and latency measures.

TABLE 4.1

Results of Mixed ANOVA to check for interactions of Latency measures across Groups and Conditions

Source	F Value	Degrees Of Freedom	Degrees Of Freedom (Error)	Significance
Condition	115.146	1	28	0.000
Groups	27.664	1	28	0.000
Condition & Groups	10.745	1	28	0.003
Latency & Groups	20019.337	3	84	0.000
Condition & Latency	9.087	3	84	0. 022
Condition & Latency & Groups	3.389	3	84	0.005

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

Descriptive statistics were also done to find out the mean and standard deviation of the latency measures for musicians and non-musicians in quiet and in noise.

TABLE 4.2

Mean and Standard Deviation for Peak Latencies (in milliseconds) of Non-Musicians and Musicians in Quiet

Parameter	Group	Mean	Standard Deviation
Latency of Wave V	Nonmusicians	6.2347	0.32711
	Musicians	6.0227	0.18892
Latency of Wave D	Nonmusicians	22.8847	0.72753
	Musicians	22.0133	0.34665
Latency of Wave E	Nonmusicians	31.0467	0.61351
	Musicians	30.4207	0.32679
Latency of Wave F	Nonmusicians	39.5400	0.66864
	Musicians	38.9713	0.23194

TABLE 4.3

Mean and Standard Deviation for Peak Latencies (in milliseconds) of Non-Musicians and Musicians in Noise

Parameter	Group	Mean	Standard Deviation
Latency of Wave V	Nonmusicians	7.6887	0.71050
	Musicians	7.4373	0.74795
Latency of Wave D	Nonmusicians	25.3507	1.50496
	Musicians	22.9620	0.91350
Latency of Wave E	Nonmusicians	33.8013	1.65300
	Musicians	31.9767	1.08499
Latency of Wave F	Nonmusicians	41.9693	2.05316
	Musicians	39.8953	0.63257

It is evident from the results that musicians showed earlier mean latencies of all the waves than non-musicians in quiet (TABLE 4.2) and in noise (TABLE 4.3).

TABLE 4.4

Paired t-test results for Latencies of waves of Non-Musicians in Quiet and in Noise

Parameter	ǀ t ǀ	Degrees of Freedom	Sig. (2-tailed)
Latency of Wave V	-9.909	14	<0.001
Latency of Wave D	-6.633	14	<0.001
Latency of Wave E	-6.859	14	<0.001
Latency of Wave F	-5.135	14	<0.001

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

TABLE 4.5

Paired t-test results for Latencies of waves of Musicians in Quiet and in Noise

Parameter	ǀ t ǀ	Degrees of Freedom	Sig. (2-tailed)
Latency of Wave V	-8.006	14	<0.001
Latency of Wave D	-3.938	14	<0.001
Latency of Wave E	-5.121	14	<0.001
Latency of Wave F	-7.371	14	<0.001

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

Paired t-test was also carried out to compare the latencies of the waves acquired in quiet and noise in each group. In both non-musicians (TABLE 4.4) and musicians (TABLE 4.5), the latencies of all the waves were seen to be significantly different in quiet and noise, with delay in latencies of the waves acquired in noise

The above findings are in agreement with Russo, Nicol, Musacchia and Kraus (2004) who documented the detrimental effects of noise on the subcortical representation of speech signals. The same findings were also reported by Parbery-Clark, Skoe and Kraus (2009).

TABLE 4.6

One Way MANOVA for Latencies of Waves in Quiet

Parameter	F Value	Degrees of freedom	Degrees of freedom (error)	Level of significance (p)
Latency of peak V	4.725	1	28	0.038
Latency of peak D	17.535	1	28	<0.001
Latency of peak E	12.165	1	28	0.002
Latency of peak F	9.684	1	28	0.004

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

TABLE 4.7

One Way MANOVA for Latencies of Waves in Noise

Parameter	F Value	Degrees of freedom	Degrees of freedom (error)	Level of significance (p)
Latency of peak V	0.890	1	28	0.353
Latency of peak D	27.614	1	28	<0.001
Latency of peak E	12.774	1	28	0.001
Latency of peak F	13.979	1	28	0.001

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

Further, One Way MANOVA tests were carried out to compare how the latencies of Waves V, D, E and F varied across the groups in quiet and in noise.

The results of the one way One Way MANOVA show that in quiet, the latencies of all the waves were found to be significantly different across the groups (TABLE 4.6). This is in agreement with the findings of Musacchia, Sams, Skoe and Kraus (2007) who found that musicians had earlier wave latencies than non-musicians in quiet. Musacchia, Stait and Kraus (2008) also documented the onset timing of musicians in quiet to be earlier than that of non-musicians. However, in contradiction Parbery-Clark, Skoe and Kraus (2009) found that the latencies of the waves were not significantly different in musicians and non-musicians in quiet.

However, the results of One Way MANOVA indicate that in the presence of noise, the latency of wave V did not differ significantly across the groups but the latencies of the D, E and F did (TABLE 4.7). In studies by Cunningham, Nicol, Zecker, Bradlow and Kraus (2001), Russo, Nicol, Musacchia and Kraus (2004) and Parbery-Clark, Skoe and Kraus (2009), it has been noted that the latency of the onset peak and transition peaks are significantly more prolonged in the presence of noise in non-musicians as compared to musicians. However, it may be noted that in the present study, the mean latencies of all the waves, including wave V are found to be earlier in musicians than in non-musicians (TABLE 4.3). In agreement with this finding, Parbery-Clark, Skoe and Kraus (2009) had found that in noise, the onset and transition responses occurred significantly earlier in musicians than in non-musicians.

These findings indicate that long term musical training not only improves the overall encoding of temporal events of the stimuli but also restricts the detrimental effects of background noise on this process (Don and Eggermont, 1978; Cunningham et al., 2001; Russo et al., 2004). The physiological basis of this finding may lie in the Medial Olivocochlear Bundle (MOCB) via which Higher-level auditory structures influence processing in lower-level structures. An increase in MOCB activity has been correlated with good speech in noise performance (De Boer & Thorton, 2008). It is possible that top-down modulation improves signal quality at the auditory periphery by selectively amplifying relevant features of the signal, and inhibiting irrelevant features in the presence of background noise. 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 top-down modulation has indeed been noted to be prominent in musicians (Trainor, Shahin, & Roberts, 2009) and an increase in top down modulation was been noted in children following a year musical training (Shahin, Roberts, Chau, Trainor, & Miller, 2008), thus indicating the role of musical training in the sharpening of the brainstem responses in noise.

Comparison of Amplitude Measures

The amplitudes of the onset peak V and the transition peaks D, E and F were considered for analysis. The amplitudes of the peaks are related to the robustness of the representation of the features of the stimulus (the onset and transition portions). The addition of background noise had been documented to result in reduction of amplitudes of the peaks of ABR, indicating a disruption in timing of brainstem activity (Don and Eggermont, 1978; Cunningham et al., 2001; Russo et al., 2004). It has been hypothesized that the disruptive effects of noise on the representation of stimulus features may be limited by long-term musical training which can bring about enhancements of stimulus features at the subcortical level via top down influences (Dean, Harper and McAlpine, 2005) mediated through the efferent auditory system (Luo, Wang, Kashani, & Yan, 2008).

Mixed ANOVA was carried out to compare between the groups for amplitude measures obtained the conditions of quiet and in noise (TABLE 4.8). The results of the test indicated as to the presence of main effects of condition and groups as well as the presence of interaction effects between conditions, groups and amplitude measures.

TABLE 4.8

Results of Mixed ANOVA to check for interactions of Latency measures across Groups and Conditions

Source	F Value	Degrees Of Freedom	Degrees Of Freedom (Error)	Significance
Condition	576.733	1	28	<0.001
Groups	19.332	1	28	<0.001
Condition & Groups	14.248	1	28	.001
Amplitude & Groups	24.940	3	84	<0.001
Condition & Amplitude	9.969	3	84	<0.001
Condition & Amplitude & Groups	53.356	3	84	<0.001

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

Descriptive statistics were done to find out the mean and standard deviation of the latency measures for musicians and non-musicians in quiet and in noise.

TABLE 4.9

Mean and Standard Deviation for Peak Amplitudes(in microvolts)of Non-Musicians and Musicians in Quiet

Parameter	Group	Mean	Standard Deviation
Amplitude of peak V	Nonmusicians	0.2247	0.04984
	Musicians	0.2773	0.05092
Amplitude of peak D	Nonmusicians	0.1573	0.07392
	Musicians	0.1880	0.03406
Amplitude of peak E	Nonmusicians	0.2260	0.09164
	Musicians	0.4147	0.09716
Amplitude of peak F	Nonmusicians	0.1600	0.049281
	Musicians	0.1731	0.072985

The examination of the mean amplitudes of the waves V, D, E and F reveals that the musicians had higher mean amplitudes than non- musicians for all the waves in quiet (TABLE 4.09). This finding is in agreement with those of Musacchia, Sams, Skoe and Kraus (2007) and Parbery-Clark, Skoe and Kraus (2009).

TABLE 4.10

Mean and Standard Deviation for Peak Amplitudes (in microvolts) of Non-Musicians and Musicians in Noise

Parameter	Group	Mean	Standard Deviation
Amplitude of peak V	Nonmusicians	0.0433	0.03177
	Musicians	0.0787	0.02326
Amplitude of peak D	Nonmusicians	0.0507	0.02549
	Musicians	0.0673	0.01534
Amplitude of peak E	Nonmusicians	0.0440	0.02746
	Musicians	0.0567	0.01496
Amplitude of peak F	Nonmusicians	0.0873	0.05378
	Musicians	0.1053	0.03543

The mean amplitudes of all the waves acquired in noise were also greater for musicians than for non-musicians (TABLE 4.10), a finding reflected in the study by Parbery-Clark, Skoe and Kraus (2009).

TABLE 4.11

Paired t-test results for Amplitudes of waves of Non-Musicians in Quiet and in Noise

Parameter	ǀ t ǀ	Degrees of Freedom	Sig. (2-tailed)
Amplitude of Wave V	10.505	14	<0.001
Amplitude of Wave D	5.922	14	<0.001
Amplitude of Wave E	7.388	14	<0.001
Amplitude of Wave F	7.542	14	<0.001

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

TABLE 4.12

Paired t-test results for Amplitudes of waves of Musicians in Quiet and in Noise

Parameter	ǀ t ǀ	Degrees of Freedom	Sig. (2-tailed)
Amplitude of Wave V	14.086	14	<0.001
Amplitude of Wave D	15.282	14	<0.001
Amplitude of Wave E	14.826	14	<0.001
Amplitude of Wave F	3.967	14	<0.001

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

Paired t-test was also carried out to compare the amplitudes of the waves acquired in quiet and noise in each group. In both non-musicians (TABLE 4.9) and musicians (TABLE 4.10), the amplitudes of all the waves were seen to be significantly greater in quiet than in noise. This indicates that the noise has a detrimental effect on the subcortical representation of the signal (Don and Eggermont, 1978; Cunningham et al., 2001; Russo et al., 2004). Russo, Nicol, Musacchia and Kraus (2004) and Parbery-Clark, Skoe and Kraus (2009) have also documented reduced amplitude of the onset and transition waves in the presence of background noise.

To compare the amplitude measures of the different waves across the 2 groups in quiet and in noise, two measures of one-way ONE WAY MANOVA were carried out.

TABLE 4.13

One Way MANOVA for Amplitudes of Waves in Quiet

Parameter	F Value	Degrees of freedom	Degrees of freedom (error)	Level of significance (p)
Amplitude of peak V	8.196	1	28	0.008
Amplitude of peak D	2.130	1	28	0.156
Amplitude of peak E	29.932	1	28	<0.001
Amplitude of peak F	0.330	1	28	0.570

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

TABLE 4.14

One Way MANOVA for Amplitudes of Waves in Noise

Parameter	F Value	Degrees of freedom	Degrees of freedom (error)	Level of significance (p)
Amplitude of peak V	12.078	1	28	0.002
Amplitude of peak D	4.709	1	28	0.039
Amplitude of peak E	2.461	1	28	0.128
Amplitude of peak F	1.172	1	28	0.288

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

The results of the one way One Way MANOVA tests show that in quiet (TABLE 4.13), the amplitudes of the waves V and E are significantly different across the groups. Parbery-Clark, Skoe and Kraus (2009) had documented that there were no significant differences in the amplitudes of the onset and transition waves in quiet across musicians and non-musicians, though the mean amplitudes were found to be greater for musicians.

In noise (TABLE 4.14), the amplitudes of waves V and D were found to be significantly different across the groups. Parbery-Clark, Skoe and Kraus (2009) also documented the reduction in amplitude of the onset and transition peaks in the presence of background noise to be similar in musicians and non-musicians. Though the mean amplitude of the transition wave was found to be greater in musicians, the same had not been observed with the onset wave. However, it may be pointed out that the amplitudes of onset responses are highly variable (Starr and Don, 1988; Hood, 1998) and this fact may have contributed to the differences present between the two studies.

From the above results, it is seen that the musicians have overall higher mean amplitudes of different waves in both quiet and in noise when compared to non-musicians. This is due to the disruption of the neural representation of stimulus features by noise (Russo et al., 2004).

However, due to the training musicians undergo which involves the selective attention to a specific element from a complex soundscape, there is an enhanced encoding which improves the subcortical signal quality, resulting in a more robust representation of the target acoustic signal in noise.

This once again points to the fact that musical training helps strengthen the sub-cortical representation of the stimulus features via top-down processes.

Comparison of Formant Amplitude Measures

The Speech ABRs acquired from the subjects in quiet and in noise were subject to Fast Fourier Transform to obtain the amplitudes of the formants of the encoded stimulus /da/. The amplitudes of the fundamental frequency (F0), which is important for identifying the speaker, and emotional tone of voice, the first formant (F1), which provides phonetic information and the second formant (F2) were considered for analysis. It was hypothesized that that the addition of noise would result in lower formant amplitudes in the presence of noise, indicating a degradation in the neural representation of the signal.

Mixed ANOVA was carried out to compare between the groups for formant amplitude measures obtained the conditions of quiet and in noise.

The results of the test are shown in TABLE 4.15.

The results indicated as to the presence of a main effect of condition while no main effect of group was noted

Interaction effects were present between formant amplitudes & groups as well as conditions & formant amplitudes while no significant effects of condition and group, condition, formant amplitude and group were noticed (TABLE 4.15).

TABLE 4.15

Results of Mixed ANOVA to check for interactions of Formant Amplitude measures across Groups and Conditions.

Source	F Value	Degrees Of Freedom	Degrees Of Freedom(Error)	Significance
Condition	82.890	1	28	<0.001
Groups	3.328	1	28	.079
Condition & Groups	0.017	1	28	.898
Formant Amplitude & Groups	3.723	2	56	.030
Condition & Formant Amplitude	64.899	2	56	<0.001
Condition, Formant Amplitude & Groups	.061	2	56	.941

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

Descriptive statistics were carried out to find out the mean and standard deviation of the formant amplitude measures for musicians and non-musicians in quiet and in noise.

TABLE 4.16

Mean and Standard Deviation for Formant Amplitudes(in microvolts) of Non-Musicians and Musicians in Quiet

Parameter	Group	Mean	Standard Deviation
Amplitude of F0	Nonmusicians	5.10913	1.645812
	Musicians	6.16387	2.059548
Amplitude of F1	Nonmusicians	0.62560	0.205510
	Musicians	0.65633	0.241572
Amplitude of F2	Nonmusicians	0.19213	0.049957
	Musicians	0.19887	0.063682

TABLE 4.17

Mean and Standard Deviation for Formant Amplitudes (in microvolts) of Non-Musicians and Musicians in Noise

Parameter	Group	Mean	Standard Deviation
Amplitude of F0	Nonmusicians	2.23173	1.315477
	Musicians	3.14567	1.738376
Amplitude of F1	Nonmusicians	0.25080	0.111335
	Musicians	0.33133	0.231192
Amplitude of F2	Nonmusicians	0.17593	0.085855
	Musicians	0.17907	0.059528

The mean values of the amplitudes of formants F0, F1 and F2 were found to be greater for musicians than non-musicians in quiet (TABLE 4.16). This is in agreement with the findings of Musacchia, Sams, Skoe and Kraus (2007). Musacchia, Stait and Kraus (2008) also found higher F0 peak amplitude in musicians than in non-musicians in quiet. However, Parbery-Clark, Skoe and Kraus (2009) found no significant difference between the formant amplitude of F0 across musicians and non-musicians in quiet though musicians did show higher mean F0 amplitude.

The mean formant amplitudes were higher for musicians than non-musicians in waveforms acquired in noise (TABLE 4.17). Parbery-Clark, Skoe and Kraus (2009) also found higher mean amplitudes of F0 in musicians than in non-musicians in noise, though the differences were not found to be statistically significant.

Paired t-test was also carried out to compare the amplitudes of the formants in quiet and noise in each group.

TABLE 4.18

Paired t-test results for Formant Amplitudes of waves of Non-Musicians in Quiet and in Noise

Parameter	$|t|$	Degrees of Freedom	Sig. (2-tailed)
Amplitude of F0	6.943	14	<0.001
Amplitude of F1	6.872	14	<0.001
Amplitude of F2	0.684	14	<0.001

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

TABLE 4.19

Paired t-test results for Formant Amplitudes of waves of Musicians in Quiet and in Noise

Parameter	ǀ t ǀ	Degrees of Freedom	Sig. (2-tailed)
Amplitude of F0	5.386	14	<0.001
Amplitude of F1	5.111	14	<0.001
Amplitude of F2	1.128	14	<0.001

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

In both non-musicians (TABLE 4.18) and musicians (TABLE 4.19), the amplitudes of all the formants were seen to be significantly different in quiet and noise. The mean values of the formant amplitudes were seen to be lesser in noise than in quiet. This is in line with the findings of Russo, Nicol, Musacchia and Kraus (2004) and Parbery-Clark, Skoe and Kraus (2009) who attributed it to the detrimental effects of noise on the neural encoding of the various formants.

To compare the amplitude measures of the different waves across the 2 groups in quiet and in noise, two measures of one-way One Way MANOVA were carried out.

TABLE 4.20

One Way MANOVA for Formant Amplitudes of Waves in Quiet

Parameter	F Value	Degrees of freedom	Degrees of freedom (error)	Level of significance (p)
Amplitude of F0	2.401	1	28	0.132
Amplitude of F1	0.141	1	28	0.710
Amplitude of F2	0.104	1	28	0.750

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

TABLE 4.21

One Way MANOVA for Formant Amplitudes of Waves in Noise

Parameter	F Value	Degrees of freedom	Degrees of freedom (error)	Level of significance (p)
Amplitude of F0	2.636	1	28	0.116
Amplitude of F1	1.477	1	28	0.234
Amplitude of F2	0.013	1	28	0.908

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

The results of the one way ONE WAY MANOVA tests show that the formant amplitudes were not found to be significantly different across the two groups in either quiet (TABLE 4.20) or in noise (TABLE 4.21). This is in accordance with the findings of Parbery-Clark, Skoe and Kraus (2009). However, Musacchia, Sams, Skoe and Kraus (2007) have documented the presence of a statistically significant difference in F0 amplitude in quiet across the two groups, with musicians showing higher F0 amplitudes than their non-musically trained counterparts, though the same findings were not true of higher formants. However, it may be pointed out that in this study, musicians did show higher mean amplitudes of all formants as compared to non-musicians.

From the above findings, it was seen that both groups also showed higher mean formant amplitudes in quiet than in the presence of noise, evidence to the degradation of the neural representation of the speech signal in the presence of noise. The musicians also showed higher mean formant amplitudes than the non-musicians in both quiet and in noise, though the differences were not statistically significant. The enhanced encoding of the formants of the speech stimulus in musicians has been documented by many authors (Musacchia et al. (2007); Wong, et al., (2007). The higher mean formant amplitudes of musicians in noise as compared to non-musicians indicates a more robust sub-cortical representation of the speech signal, possibly brought about by years of continuous musical training. One possible explanation for this finding is the based on the 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). Given the present results, we can speculate that extensive musical training may lead to greater neural coherence, especially pertaining to relevant features crucial to the identification of the stimulus. This strengthening of the underlying neural circuitry would lead to a better bottom-up, feed-forward representation of the signal. We can also interpret these data within the framework of corticofugal modulation in which cortical processes shape the afferent auditory encoding via top-down processes as mentioned earlier in the discussion. Though we cannot separate the contributions of top-down and bottom-up processing, they are not mutually exclusive explanations. In all likelihood, top-down and bottom-up processes are reciprocally interactive with both contributing to the subcortical changes observed with musical training.

Comparison of SPIN Scores

The performance of the subjects on the task of speech perception in noise was measured in terms of percentage correct scores on the SPIN test which used the Phonemically Balanced Wordlist in Kannada (Yathiraj and Vijayalakshmi, 2005) presented at 0dB SNR in a background of speech noise. It was speculated that the disruption of neural timing and encoding of stimulus features in the presence of competing noise would be lesser in musicians than in non-musicians, resulting in enhanced performance on the task of speech perception in noise.

An independent sample t-test was carried out to compare SPIN scores across the 2 groups. The results indicate that the scores differ significantly across the 2 groups (TABLE 4.22).

TABLE 4.22

Results of Independent Sample t-Test for comparing SPIN Scores across the 2 groups

Parameter	ǀ t ǀ	Degrees of Freedom	Sig. (2-tailed)
SPIN scores	-3.500	28	0.002

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

It may be seen from TABLE 4.23 that the musicians had a higher mean score on the SPIN test than the non-musicians. Parbery-Clark, Skoe, Lam and Kraus (2009) and Parbery-Clark, Skoe and Kraus (2009) also report of a distinct advantage in musicians on the task of perception of speech in noise. This advantage was reported to correlate well with the number of years of training the musician had undergone, which strongly suggested that such intensive training helps to fine tune sensory and cognitive processes that contributed to the task of speech perception in noise.

TABLE 4.23

Mean and Standard Deviation of SPIN Score In Musicians and Non-Musicians

Group	Mean (in percentage)	Standard Deviation (in percentage)
Non- Musicians	78.667	3.266
Musicians	82.400	2.529

To investigate whether the superior performance of musicians over non-musicians on the task of speech perception in noise was related to the differences in the subcortical encoding of speech stimuli across the two groups, Pearson Correlation Coefficient was calculated to check whether the SPIN scores correlated to the different latency (latencies of waves V, D, E and F), amplitude (amplitudes of waves V, D, E and F) and formant amplitude (formant amplitudes of F0, F1 and F2) measures in quiet and in noise.

In non-musicians, the SPIN scores did not correlate with any of the measures obtained in quiet (TABLE 4.24).

In musicians, the SPIN scores were found to correlate negatively with the latencies of waves D and F obtained in quiet (TABLE 4.26). Parbery-Clark, Skoe and Kraus (2009) found that in quiet, there was no significant correlation between latency, amplitude or formant amplitude of brainstem responses of a subject and the corresponding scores on the task of speech perception in noise.

TABLE 4.24

Results of Pearson's Correlation: Correlation of SPIN Scores with Latency, Amplitude and Formant Amplitude in Quiet for Non-Musicians

Parameter	Pearson Correlation	Significance (2-tailed)
Latency of Wave V	-0.478	0.072
Latency of Wave D	0.000	0.998
Latency of Wave E	-0.084	0.767
Latency of Wave F	0.201	0.471
Amplitude of Wave V	-0.170	0.545
Amplitude of Wave D	-0.241	0.388
Amplitude of Wave E	0.334	0.224
Amplitude of Wave F	-0.036	0.900
Amplitude of F0	-0.062	0.826
Amplitude of F1	-0.029	0.920
Amplitude of F2	-0.013	0.964

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

TABLE 4.25

Results of Pearson's Correlation: Correlation of SPIN Scores with Latency, Amplitude and Formant Amplitude in Quiet for Musicians

Parameter	Pearson Correlation	Significance (2-tailed)
Latency of Wave V	-0.493	0.062
Latency of Wave D	-0.655	0.008
Latency of Wave E	-0.334	0.224
Latency of Wave F	-0.610	0.016
Amplitude of Wave V	0.231	0.408
Amplitude of Wave D	-0.239	0.391
Amplitude of Wave E	0.451	0.092
Amplitude of Wave F	0.158	0.573
Amplitude of F0	0.195	0.487
Amplitude of F1	0.130	0.644
Amplitude of F2	-0.250	0.369

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

In noise, it was seen that for non-musicians, the SPIN scores negatively correlated with the latency of wave V and positively correlated with the amplitudes of wave V and

D (TABLE 4.26). Thus, poorer performance on the SPIN test was found to be related to the prolongation of onset latency and the reduction of amplitudes of the onset wave V and transition wave D in non-musicians, indicating that addition of noise had resulted in disruption of brainstem timing and a reduction in the amplitude of the responses encoding stimulus features (onset and transition), which had resulted in reduced SPIN scores.

In noise, the SPIN scores of musicians correlated negatively with the latencies of wave V and F (TABLE 4.27), indicating that subjects with earlier wave V and F latencies showed better performance on the task of speech perception in noise. Positive correlation was seen with the formant amplitude of F0 obtained in noise, indicating that the superior encoding of F0 in musicians had resulted in enhanced SPIN scores.

These findings indicate that musical training results in an increased resistance of the brainstem response to the disruptive effects of background noise, resulting in better timing of brainstem responses and the better encoding of stimulus features.

TABLE 4.26

Results of Pearson's Correlation: Correlation of SPIN Scores with Latency, Amplitude and Formant Amplitude in Noise for Non-Musicians

Parameter	Pearson Correlation	Significance (2-tailed)
Latency of Wave V	-0.788	<0.001
Latency of Wave D	0.070	0.804
Latency of Wave E	0.015	0.959
Latency of Wave F	0.296	0.284
Amplitude of Wave V	0.541	0.037
Amplitude of Wave D	0.561	0.030
Amplitude of Wave E	-0.159	0.571
Amplitude of Wave F	0.027	0.924
Amplitude of F0	0.164	0.559
Amplitude of F1	0.284	0.304
Amplitude of F2	-0.081	0.775

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

TABLE 4.27

Results of Pearson's Correlation: Correlation of SPIN Scores with Latency, Amplitude and Formant Amplitude in Noise for Musicians

Parameter	Pearson Correlation	Significance (2-tailed)
Latency of Wave V	-0.886	<0.001
Latency of Wave D	-0.388	0.153
Latency of Wave E	-0.096	0.734
Latency of Wave F	-0.692	0.004
Amplitude of Wave V	0.010	0.973
Amplitude of Wave D	0.250	0.368
Amplitude of Wave E	0.453	0.090
Amplitude of Wave F	0.006	0.982
Amplitude of F0	0.517	0.048
Amplitude of F1	0.378	0.165
Amplitude of F2	0.338	0.217

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

The findings in noise are in line with those of Parbery-Clark, Skoe and Kraus (2009) who also documented a correlation between better scores on the HINT and earlier latencies of onset and transition waves. However, the same study did not document a correlation with F0 amplitude as was seen in the present study. This may be because of the difference in the maskers used during the test of speech perception in noise. While in the present study, speech noise had been used, Parbery-Clark, Skoe and Kraus (2009) had used multi-talker babble which is a more realistic approximation of competing signals one might encounter in real life.

The higher mean SPIN scores of musicians as compared to non-musicians indicate that they have a superior ability to detect speech signals in a background of competing noise. This is a consequence of their intensive training that render them experts in extracting relevant signals from complex soundscapes. A distinct advantage is seen in musicians on the task of perception of speech in noise, which correlated strongly with the number of years of consistent practice (Parbery-Clark, Skoe, Lam and Kraus, 2009). Musical experience was seen to resulted in more robust sub-cortical representation of speech in the presence of background noise, which may contribute to musician’s behavioral advantage for speech in noise perception (Parbery-Clark, Skoe, and Kraus, 2009). Musicians also exhibited more faithful encoding the steady state portion of a stimulus in the presence of background noise and had higher stimulus-to-response correlations in noise than non-musicians which is indicative of more precise neural transcription of stimulus features. These enhancements may be related to the effects of the top-down (Suga, Zhang and Yan, 1997; Zhang, Suga and Yan, 1997; Luo, Wang, Kashani & Yan, 2008) and bottom-up processes (Hebb, 1949) that act to reduce the disruptive effects of noise while selectively enhancing stimulus features. These enhancements mean that the important features that contribute to speech intelligibility are still represented faithfully at the level of the brainstem despite the presence of a disruptive background noise. This would translate into an improved perception of speech in the presence of a competing signal.

Thus this study indicates that listening and training experiences of musicians modulate their neural responses in such a manner as to allow for enhanced perception of speech stimuli in competing backgrounds.

CHAPTER 5

SUMMARY AND CONCLUSION

The present study was carried out with the aim of verifying whether Carnatic musicians had a more robust encoding of speech stimuli at the level of the brainstem as compared to non-musicians in the presence of a competing noise and whether such an enhancement would result in an enhanced ability to perceive speech in noise.

The speech ABR was recorded in quiet and in the presence of noise (0 dB SNR) using the /da/ stimulus. The latencies and amplitudes of the onset peak V and the transition peaks D, E and F were noted and the waveforms were analyzed using FFR to obtain the amplitudes of the encoded formants F0, F1 and F2.

The results revealed that musicians had a more robust representation of the speech signal at the level of the brainstem as evidenced by earlier latencies, higher amplitudes and higher amplitudes of encoded formants than in non-musicians.

The presence of noise was also found to affect musicians less than non-musicians, with better morphologically defined waveforms, earlier latencies, higher amplitudes and higher amplitudes of encoded formants than in non-musicians.

These findings were in agreement with a number of studies that preceded it. It adds to the growing literature on the effect on musical training on neural plasticity and the neural representation of signals at the level of the brainstem.

FUTURE RESEARCH CONSIDERATIONS

Though the results of the present study provide biological evidence for the positive effect of lifelong musical training on speech-in-noise encoding, we cannot determine the extent to which this enhancement is mediated directly by musical training, group genetic differences, or a combination of the two. Longitudinal studies, akin to the large-scale design recently described by Forgeard et al. (2008) and Hyde et al. (2009), could not only elucidate the course of developmental of these skills and/or the genetic disposition for the musician neural advantage for speech-in-noise, but it may also help disentangle the relative influences of top-down and bottom-up processes on the neural encoding of speech-in-noise.

Other important lines of research include the impact that the choice of musical instrument and musical genre, as well as extensive musical listening experience in the absence of active playing, have on the subcortical encoding of speech-in-noise..

By studying a population that displays enhanced neural speech encoding, we can investigate which factors contribute to an enhanced ability for speech perception in noise, providing future avenues for the investigation of speech perception deficits in noise as experienced by older adults and hearing-impaired and language-impaired children. By providing an objective biological index of speech perception in noise, brainstem activity may be a useful measure for evaluating the effectiveness of Speech-In-Noise based auditory training programs.

REFERENCES

Abdala, C., & Folsom, R. C. (1995). Frequency contribution to the click-evoked auditory brain stem response in human adults and infants. *Journal of the Acoustical Society of America*, *97*(4), 2394–2404.

Anderson, S., Skoe, E., Chandrasekaran, B., Zecker, S., & Kraus, N. (2010). Neural signatures of speech-in-noise and reading in the auditory brainstem. Manuscript submitted for publication.

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.

Bent, T., Bradlow, A. R., and Wright, B. A. (2006). The influence of linguistic experience on the cognitive processing of pitch in speech and non-speech sounds. *Journal of Experimental Psychology, 32*, 1, 97–103.

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.

Bengtsson, S.L., Nagy, Z., Skare, S., Forsman, L., Forssberg, H., Ullén, F. (2005). Extensive piano practicing has regionally specific effects on white matter development. *Nature and Neuroscience 8 (9),* 1148–1150.

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

Bidelman, G. M., Gandour, J. T., & Krishnan, A (in press). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience*.

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

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.

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 from event related potential recordings. *Music Percept*, *19*, 199–222.

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., 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.

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.

Cunningham, J., Nicol, T., Zecker, S., & Kraus, N. (2000). Speech-evoked neurophysiologic responses in children with learning problems: development and behavioral correlates of perception. *Ear and Hearing*, *21*, 554-568.

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

Dean I, Harper NS, McAlpine D (2005) Neural population coding of sound level adapts to stimulus statistics. *Nature and Neuroscience, 8,*1684–1689.

De Boer, J., & Thornton, A. R. (2008). Neural correlates of perceptual learning in the auditory brainstem: Efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. *Journal of Neuroscience*, *28,* 4929-4937.

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., (2004). Musical Training Enhances Automatic Encoding of Melodic Contour and Interval Structure(2004*). Journal of cognitive Neuroscience 16,* 1010-1021.

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

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

Gorga, M.P., Kaminski, J.R., Beauchaine, K.L.,Jesteadt, W., & Neely, S.T. (1989) Auditory brainstem responses from children three months to three years of age: normal patterns of response. *InternationaI Journal of Speech and Hearing Research 32*:281–288.

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

Helfer, K.S. & Vargo, M. (2009) Speech recognition and temporal processing in middle-aged women. *Journal of Acoustical Society of America, 20(4)*:264 71.

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

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*, 393-402.

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.

Hornickel, J., Skoe, E., Nicol, T., Zecker, S., & Kraus, N. (2009). Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 13022-13027.

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. *Psychol Music*, *25*, 70–83.

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., & Gandour, J. T. (2009). The role of the auditory brainstem in processing linguistically relevant pitch patterns. *Brain and Language*, *110* , 135-148.

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.

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

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

Luo, F.,Wang, Q., Kashani, A., & Yan, J. (2008). Corticofugal modulation of initial sound processing in the brain. *Journal of Neuroscience*, *28,* 11615-11621.

Magne, C., Schon, D., & Besson, M. (2006).Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, *18,* 199-211.

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.

Micheyl, C., Delhommeau, K., Perrot, X., and Oxenham, A. J., (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hear Res, 219*, 36–47.

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.

Musacchia, G., Strait, D., & Kraus, N., (2008). Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians, *Hearing Research 107,* 15666-15670

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.

Ohishi, 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., & 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.

Perrot, X., Ryvlin, P., Isnard, J., Guenot, M., Catenoix, H., Fischer, C., et al. (2006). Evidence for corticofugal modulation of peripheral auditory activity in humans. *Cerebral Cortex*, *16,* 941-948.

Ponton, C.W., Eggermont, J.J., Coupland, S.G. & Winkelaar, R. (1992) Frequency-specific maturation of the eighth nerve and brain-stem auditory pathway: evidence from derived auditory brain-stem responses (ABRs). *Journal of the Acoustical Society of America, 91*:1576–1586.

Rammsayer, T., & Altenmuller, E. (2006). Temporal information processing in musicians and nonmusicians. *Music Percept*, *24*, 37–48.

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.

Salamy, A. (1984) Maturation of the auditory brainstem response from birth through early childhood. Journal of Clinical Neurophysiology *1*:293–329.

Schlaug G, Jancke L, Huang Y, Staiger JF, Steinmetz H (1995a) Increased corpus callosum size in musicians. *Neuropsychologia 33*: 1047-1055.

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., Specht, H.J., Gutschalk, A., & Rupp,A. (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.

Skoe, E. & Kraus N. (2010) Hearing it again and again: on-line subcortical plasticity in humans. *PlosONE*. 5(10)

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 Masters dissertation submitted to University of Mysore.

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

Shahin , A., Roberts , J., & Trainor, L., (2004). Enhancement of auditory cortical development by musical experience in children. *Auditory and Vestibular Systems, 15*, 1917-1921.

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.

Tervaniemi, M., Kruck, S., Baene, W. D., Schröger, E., Alter, K., & Friederici, A. D. (2009). Top-down modulation of auditory processing: Effects of sound context, musical expertise and attentional focus. *European Journal of Neuroscience, 30,* 1636-1642.

Thomas, A.O., (2011). Unpublished Masters dissertation submitted to University of Mysore.

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., (2009). 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.

Ziegler, J. C., Pech-Georgel, C., George, F., & Lorenzi, C. (2009). Speech-perception-in-noise deficits in dyslexia. *Developmental Science, 12*, 732-745.

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.