PERIODICITY CODING IN CHILDREN WHO ARE LEARNING INSTRUMENTAL CARNATIC MUSIC

Manjula C

Student Register Number: 13AUD012

This Dissertation is submitted as part fulfillment
for the Degree of Master of Science in Audiology
University of Mysore, Mysore

ALL INDIA INSTITUTE OF SPEECH AND HEARING
MANASAGANGOTHRI
MYSORE 570 006

May, 2015

Certificate

This is to certify that this dissertation entitled **“Periodicity coding in children who are learning instrumental Carnatic music”** is a bonafide work in part fulfillment for the degree of Master of Science (Audiology) of the student with Registration No. 13AUD012. 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.

Mysore,
May 2015

Prof. S. R. Savithri
Director
All India Institute of Speech and Hearing
Manasagangothri,
Mysore – 570006

Certificate

This is to certify that this dissertation entitled **"Periodicity coding in children who are learning instrumental Carnatic music"** has been prepared under my supervision and guidance. It is also certified this has not been submitted earlier in other University for the award of any other Diploma or Degree.

Mysore,
May, 2015

Dr. K. Rajalakshmi
Guide
Professor of Audiology
Department of Audiology
All India Institute of. Speech and Hearing
Manasagangothri, Mysore – 570006

Declaration

This dissertation entitled **Periodicity coding in children who are learning instrumental Carnatic music** is the result of my own study under the guidance of Dr. K. Rajalakshmi, Professor in Audiology, Department of Audiology, All India Institute of Speech and Hearing, Mysore, and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysore, **Register No**: 13AUD012

May, 2015

Dedicated to
OUR Beloved GUIDE
Dr. RAJALAKSHMI
APPA
AMMA
&
best forever SUMAN BHAJUUU

Acknowledgement

I would like to thank the god almighty, who made me to complete the dissertation on time...

I would like to thank Director madam, **Dr. S.R Savitri** for giving the opportunity to do the dissertation....

I would like to thank the HOD of audiology **Dr Ajith Kumar U** for his constant support, suggestions and guidance during the dissertation...

I thank my beloved guide **Dr. K Rajalakshmi** for her valuable suggestions and encouragement for the dissertation. Thank you very much ma'am for your concern.

I am very greatful to **Vasanthalakshmi mam...** for timely help and guidance.....

Without thanking tooo many, my dissertation won't be complete....

My sincere thanks to the participants of the study...

A big thanks to all the Staffs of Dept of Audiology....

Spoorthi akka and Sumanth... thanks a lot for ur good and valuable suggestions through out the dissertation....

Heartful thanks to **Nirmala, Sneha, Navya, Potha sounder Raj, Pallavi, harshita, sanjana, sujana** and many more...

Thanks for the support machis.... **Harika, Kavya, padma, navya, suman, zeena, vineetha, amritha,.....**

Heartful thanks to **soulmate Kiranaaa and bebek....** For their timely help & support...

A special thanks to **JSS PUBLIC SCHOOL, SJCE** campus, Manasagangothri... and **Subbanna Keshava Music School,** vidyaranyapuram.....

Special thanks to **Dee, Ko, Ashok Bhava, Manju Bhava, uncle and aunt.... Heartful thnks to Manaswi ashok, Jeevan, Tandav, Cheluvaraj....**

Without your support this would not be possible... **thanku Raaammm...**

Thanks **Neelesh Benet and Suman,** for being with me throughout the dissertation.... Thanks a lot.... U ppl are awesome... I am happy to have ppl like u as my batchmate and bestieee too....

I ll miss my best postings partners.... **Anjali, Neelesh, Swathi, Ayesha, Kiran, Sneha, Supreeth, Padma, amritha.......**

Thanks for **MAFIA FAMILY** members.... anjali, madhuri, nirmala, bhuvana, kiran, husna, thareeque, hasheem, akhil, amritha, navya, neelesh, akshay, bebek, himanshu, jithin ....

Thanku himanshu kr Sanju.... **Heman**... thanks for ur guidance through out....

A warm n heartful thanks to special person... for your care, corcern n guidance through out the PG life.... Miss u....

Thanks to my family members.. for their support in my life......

Thank you one and all....

Abstract

The present study aimed to investigate the neural encoding of the pitch in children who are learning Carnatic instrumental music. A total of 30 children comprising 15 children who are learning instrumental music and 15 children who have not learnt music. Behavioral measure Speech in Noise test was carried out at 0 dB SNR and electrophysiological Speech evoked auditory brainstem was recorded using 40 ms /da/ stimulus in both quiet and noise condition. Fast fourier Transform was done to see the strength of the fundamental frequency and its harmonics. Latency and amplitude of the speech evoked auditory brainstem response was also estimated in both quiet and noise condition. The results show that, F0 coding was better in instrumental musicians than in nonmusicians. The encoding of the fundamental frequency and its harmonics were also better in quiet condition than in noise condition.There was no correlation found between the Speech Perception in Noise and the amplitude of F0, F1 and F2 in quiet condtion for both musicians and nonmusicians. The latency and amplitude was robust in musicians than in nonmusicians. The enhancement seen in the instrumental musicians could be due to the fine tuning of the neural pathways and also more synchronous neural phase locking ability. Also the top-down process influencing the musicians to have greater pitch encoding.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Without music, life would be a mistake ’- Friederich Nietzsche

Music is known as a direct way to convey feelings from ancient times. Music is a complex auditory task and musicians spend years fine-tuning their skills and provide pleasure for all. Indian Music is probably the most complex musical system in the world with a very highly developed melodic and rhythmic structure. There are 2 major systems of music in Indian subcontinent. They are Hindustani and Carnatic Music. Carnatic music is the south Indian classical music . In Carnatic music it has three folds that are Vocal Music, Instrumental Music, and Dance. Vocal Carnatic musicians are those singers who sing in the concert or orchestra. Instrumental music, where notes are produced by musical instruments.

Learning to play an instrument is a complex task involving higher cognitive functions and interactions of multimodality which results in structural, functional and behavioral changes. The changes to happen takes from months to years together. There are studies which have reported that experience can modify both structural and functional changes in the auditory system. The changes seen are at the various levels of the auditory system, from the auditory brainstem to primary and surrounding auditory cortices (Wong et al., 2007). Music is an effective medium for improved auditory skills which demonstrates benefits in cognition, speech perception and others.

The neural encoding of sound begins in the auditory nerve and travels to auditory brainstem to stimulate. The encoding of speech stimulus gives insight about the central auditory pathway involved in communication. The speech evoked auditory brainstem responses can be used to study encoding of the pitch in various populations. The clinical applications of speech

ABR are to identify the auditory based learning disabilities, predicting future language impairment, predicting success with auditory training. A better understanding of neural encoding helps with early diagnosis and intervention of auditory disorders and also to measure success of training programs such as language training and musical training.

Music and language share the many features in many aspects, 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 the prosodic information; listeners use prosodic cues to identify indexical information that is, information about the speaker's intention as well as emotion and other social factors. Listening to music involves both high cognitive demands and auditory acuity; these subcortical enhancements may result from corticofugal (top–down) mechanisms.

Pitch is considered as the perceptual correlate of the fundamental frequency. From literature we can say that, musicians have more strong and accurate encoding of pitch than non-musicians (Wong, Skoe, Russ & Kraus, 2007; Clark , Strait & Kraus, 2011 and Skoe & Kraus, 2012). Apart from this there is an improvement in the capability to discriminate the pitch after musical training (Moreno, Marques, Santos & Besson, 2012).

Musicians can discriminate the tiny pitch changes than non musicians. Compared to non musicians, musicians have a variety of cortical and perceptual specializations (Musacchia, Strait & Kraus, 2008). They have the ability to discriminate and detect the minute changes in the pitch (Krishnan et al., 2010). This is because of their ability to pay more attention to detect the smaller changes in acoustic stimulus (Musacchia et al., 2007).

We know that speech evoked auditory brainstem response is extensively used to assess the speech perception abilities and provides a biological marker of an individual's auditory processing characteristics. The speech evoked auditory brainstem response consists of 2 major portions – an initial peak-trough complex, typically seen in the latencies less than 10-12 ms when evoked by a CV syllable, represents the transient portion of the stimulus associated with the burst portion of the consonant part of the syllable. Then followed by a series of peaks which represent the sustained portion of the vowel part of the syllable. The sustained portion is called as the frequency following response (FFR). The initial peak-trough complex, labelled as V-A, is thought to be analogous to the wave V of the click evoked ABR. The FFR portion, which follows the V-A complex, contains a series of peaks labelled C, D, E, F and O. The peak labelled 'O' represents the offset of the stimulus. The salient feature of the sustained portion is the periodicity, which follows the frequency information contained in the response. Hence it named as Frequency Following Response (FFR).

The FFR is often analysed using Fast Fourier Transformation (FFT) to evaluate the energy contained in the regions corresponding to the fundamental frequency of the stimulus and its harmonics. Typically, FFR is analysed to find out the strength of the F0 and the subsequent two higher formants, F1 and F2. Overall, it can be concluded that the transient portion of the brainstem response is assumed to be the neural correlate of rapid temporal changes inherent in the consonant portion, whereas, the sustained portion can be considered to encode the information related to the periodicity of the fundamental and the harmonics.

Studies conducted with the musicians using speech evoked auditory brainstem responses, had showed a lag in the neural response timing reflecting the neural delays in the auditory brainstem. Following the musical training Musacchia, Sams, Skoe and Kraus (2007) recorded speech evoked auditory brainstem response in 29 adults with normal hearing and with a background of musical training. They found that musicians had early latency compared to non musicians and also the phase locking ability was enhanced to the periodicity of the acoustic stimulus. Speech evoked F0 amplitudes correlated with the experience in musical training. The result also suggests that strength of the pitch coding strongly correlates with the intensive musical practice and exposure.

Need for the study

Musacchia et al (2007) has reported that perception of pitch was enhanced in musicians and it was strongly correlated to the years of musical experience. This study was carried out in western musicians. Rajalakshmi and Anoop (2012) reported that, reduced F_0 coding in the presence of hearing loss.

There is scarce Indian literature on the encoding of pitch in musicians. Hence, it is essential to carry out the study on encoding of pitch in children who are enrolled for instrumental Carnatic music classes.

Aim of the study

The aim of the present study is to observe the neural encoding of pitch in children who are leaning Carnatic Instrumental music.

Objectives of the study

1. To investigate the pitch encoding mechanism in children with Instrumental Carnatic musical training using the speech evoked auditory brainstem responses.
2. To observe the relation between pitch encoding mechanism and speech perception abilities in musicians.

Chapter 2

Literature Review

Structural and functional changes seen in central auditory nervous system

In the human brain training induces both changes in structural and functional aspects. There are various methods to study these structural and functional related changes seen in auditory system following musical training. It includes non invasive imaging methods such as Electroencephalography (EEG), Magnetoencephalograpy (MEG), Magnetic resonance imaging (MRI) and functional Magnetic resonance imaging (fMRI). In the electrophysiological approach, EEG and MEG show the brain's electrical and magnetic activity respectively at the scalp. Due to musical training, MRI and fMRI reveals the neuroanotomical and functional loci of the undergoing changes (Ohnishi et al 2001).

From the various studies done so far, it is understood that musicians have the nervous system distinct from the non-musicians due to the musical training. The training related changes seen at the brainstem level is recorded using Frequency Following Response (FFR). The FFR is more robust in the musicians than non-musicians (Musacchia et al 2007). At the primary auditory cortex level, Middle Latency Response (MLR) is used to trace the musical training related changes. Studies have shown that for both pure tones and music sounds, musicians have larger MLR components (Na/Pa/Nb/P1) indicates the enhanced pitch and rhythm coding relative to non-musicians.

To study the structural and functional changes seen in the human brain after musical training Hyde, Lerch, Norton, Foregeard, Winner, Evans, and Schlaug (2009) considered two groups of participants one with formal instrumental musical training and another without any formal musical training. Both behavioral tests and MRI scanning was carried out. Results suggested that in behavioral tests musicians outperformed compared to the nonmusicians. In the MRI scanning the children learning instrumental music showed significantly different brain deformations than nonmusicians. Changes seen were in the areas of greater relative voxel size in motor areas such as precentral gyrus, the corpus callosum and in the right primary auditory cortex. The results suggested that training induces structural brain plasticity early in childhood. It was concluded that long term training intervention programs in children facilitate neural plasticity.

To summarize with musical training, there is evidence of the structural and functional changes seen in the different levels of the brain.

Plasticity due to experience related training

A study was taken up by Bidelman and Alain (2015) to study the brain plasticity using the event related potential in 10 instrumental musicians and 10 nonmusicians. This study results suggest that musicians showed earlier brainstem response latencies than their non-musicians peers. This indicates that individuals with musical training showed more robust and faster neural encoding at both brainstem level and cortical level.

There is evidence that individuals with musical training are associated with greater auditory working memory capacity across the life span (Kraus et al, 2012). Musical training also leads to better performance in everyday listening tasks such as speech perception in the presence of background noise.

Strait, O’Connell, Parbery-Clark, and Kraus (2013) studied the auditory brainstem response recorded for stop consonants /ba/ and /ga/. The participants were 76 normal hearing children and adults including both musicians and non-musicians in the age range of 3-30 years.

The results demonstrated that musicians had more temporally distinct responses for both /ba/ and /ga/ and also reveled more distinct neural encoding of stop consonants arising as early as 3 years in the life.

Strait, Slater, O'Connell and Kraus (2015) studied the relationship between musical training and the development of neural mechanism of selective auditory attention. The study included 78 normal hearing participants including both adults and children. Later they divided into three age categories such as pre-schoolers, school aged children and adults. Again the subjects in each group were subdivided into musicians and non-musicians. /da/ stimulus was generated using Klatt-based synthesizer and recorded speech evoked brainstem response for the stimulus. The stimulus was presented through loud speaker. The results obtained from all age groups reflected the well-established developmental characteristics. In the groups, the characteristic maturational changes showed as earlier latencies and larger amplitudes were noticed with development. Also, it suggests that there are heightened attentional effects on auditory evoked response variability over the prefrontal cortex emerging in musicians during the early childhood. From this study it can be concluded that, incorporation of musical training into the early childhood educational programs provides the preventative effects of the behavioral disorders such as attention deficit hyperactivity disorder.

Studies have shown that music programs can impact on developing nervous system. One study done by Kraus, Slater, Thompsen, Hornickel, Strait, Nicol and White-Schwoch in 2014 to evidence how musical training improves the neural encoding of the speech. They considered 44 children with a mean age of 8.25 years. Two groups were created with group 1 having one year of musical training and group 2 with two years of musical training. Speech evoked auditory brainstem response was recorded for the stimulus /ba/ and /ga/. The results revealed that there was enhancement of neurophysiological functions with musical training progressively with the age. Also children with 2 years of musical training had improvement in the neural differentiation of syllables /ba/ and /ga/. Thus more musical training leads to larger enhancements in the neural functions, revealing that there is neuroplasticity in the impoverished human brain. Hence, the child with academic difficulties, if engaged in the musical training can bring about biological changes in the neural process which is important in the everyday communication. To conclude, if musical training is introduced during the sensitive period for auditory development, it results in improving auditory processing skills.

Speech perception abilities and musicians

A study done by white-Schwoch, Car, Anderson, Strait, and Kraus (2013), studied speech evoked auditory brainstem responses for /da/ stimulus in both quiet and noisy condition in 44 older adults in the age range of 55 to 76 years. The subjects were grouped based on the formal musical training in their childhood, such as none with 0 years of musical experience, little and moderate with 1-3 years and 4-14 years of musical training respectively. The results suggest that, adults with moderate musical training had faster neural timing in reponse to the /da/ stimulus presented in both quiet and noisy condition. The response to the /da/ stimulus is generated by the synchronous firing of the midbrain nuclei, mainly the Inferior colliculus (Chandrashekaran and Kraus, 2010). This suggests that musical training in adolescence and young adulthood may carry meaningful biological benefits into older adulthood. The musical training can offset the age related declines in cognitive and neural functions. Limited musical training in the age of 5-10 years, can lead to the neural enhancement seen after the training has stopped (Skoe and Kraus, 2012).

Parbery-Clark, Tierney, Strait, and Kraus (2012) conducted a study on musicians to study that the musicians had enhanced neural responses to music and speech stimuli including changes in the pitch, duration, intensity, timbre and voice onset time. This study was undertaken with 23 musicians and 27 non-musicians. They recorded speech evoked auditory brainstem response for three speech syllables /ba/, /da/ and /ga/ which was 170 ms in duration. A behavioral test Quick Speech Perception in Noise test was also carried out. Results of QUICKSIN test revealed that musicians performed better than non-musicians and thus better performance are due to the greater neural distinction of speech sounds which can account for advantage of hearing in adverse listening conditions. Speech ABR showed that musicians had greater neural differentiation of the 3 speech sounds. This concludes that musical training early in life provides an effective rehabilitation approach for children who have difficulty in reading and hearing in noise.

To study the musician's enhanced speech perception in noise, Parbery-Clark, Anderson, Hittner and Kraus (2012) aimed to see the effects of musical training on the subcortical responses to speech and speech perception in noise in middle aged adults. They considered 23 instrumental musicians and 17 non-musicians and recorded speech evoked ABR using /da/ stimulus of 170 ms. The results showed that, all the participants had distinct transitions and vowel peaks for all the subjects both in quiet and noise condition. FFT was done in MATLAB (version, 2009b) to estimate the neural encoding of the stimulus spectrum. Behavioral Hearing in Noise Test (HINT) was also administered. The results obtained revealed that musicians had greater speech in noise perception and suggested that they had less difficulty hearing in noise than the non-musicians. The musicians also had earlier neural responses timing, greater neural representation of the stimulus harmonics as well as more precise phase locking to the stimulus, both temporal envelope and stimulus-response correlations. The robust brainstem responses to speech correlating with the better speech in noise performance. Musicians had enhanced onset and transition timing in quiet and limited degradative effects of background noise for all aspects of neural timing. In the spectral representation of the harmonics, in both quiet and noise conditions, musicians had a robust auditory brainstem representation of the F0 and harmonics than the non-musicians. The spectral amplitude for musicians was greater than the non-musicians. The authors concluded that, because of the strengthening of the spectral features, musicians had the ability to segregate noise and advantage of speech in noise performance.

The musicians had a better neural representation of the stimulus envelope in both quiet and noise condition. In noise condition, had a significant effect on envelope encoding for both groups, envelope got stronger in noise. The representation of F0 or the harmonics are not directly related to speech in noise perception. It is because of subcortical measures associated with better speech in noise perception. From the above study it can be concluded that, musical training across the life span will have a pervasive effect on sensory and neural processing maintaining neural functions both in quiet and in noise condition.

Parbery-Clark, Skoe and Kraus (2009), carried out a study on 16 instrumental musicians with more than 10 years of musical training and 15 non-musicians with less than 3 years of musical training. 170 ms /da/ stimulus was used to record the speech ABR, and stimulus was presented at 80 dBSPL in quiet, with +10dB SNR over the background babble. Behavioral measures, HINT and Quick Speech-in-Noise (QuickSIN) was also administered. The results obtained show that musicians had more robust speech evoked auditory brainstem response in the presence of noise. Also musicians had earlier response onset timing and greater phase locking to the temporal waveform and stimulus harmonics than non-musicians. Earlier response timing was correlated with better speech in noise perception as measured through Hearing In Noise Test. With musical training, the effect of background noise on the peaks of the brainstem responses is less. Musicians exhibited smaller delays in timing than non-musicians. In the coding of the harmonics, musicians had significantly greater encoding of harmonics than non-musicians. The greater representation is associated with the highest degree of correlation between the stimulus and response. The behavioral and brainstem relationship revealed that better behavioral speech in noise perception is associated with greater precision of brainstem timing in the presence of background noise. The musicians used the fine grained acoustic information and life long 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. The enhanced encoding improves the subcortical signal quality, resulting in robust representation of target acoustic signal in noise. Top-down and bottom-up process are reciprocally interactive with both contributing to subcortical changes observed with musical training. Harmonics underlie the perception of timbre. Hence, the differentiation between two musical instruments producing same tone, harmonics are important. Musicians were more sensitive to subtle harmonic change both behaviorally and cortically. This could be interpreted as musicians had advantage of stream segregation, which is important for speech in noise perception.

Strait, Parbery-Clark, O’Connell and Kraus (2013) studied the speech evoked auditory brainstem response in 18 musicians and 14 nonmusicians using /da/ stimulus. Musicians showed faster neural response to speech onset and formant transition in both quiet and noise conditions. In the musician group, with the addition of continued musical training further, protects the neural responses from the degradative effects of the noise. Also the musicians had subcortical auditory processing advantage compared to nonmusicians (Bidelman et al, 2007; Musacchia et al, 2007).

Neural Synchrony and musicians:

Musicians have accurate neural phase locking to the temporal envelope periodicity of the stimulus. From the study we know that brainstem response represents the temporal envelope of the musical intervals. The inferior colliculus phase locks to the envelope and frequency of the amplitude modulation of the stimulus (Lee, Skoe, Kraus and Ashley, 2009).

Studies have demonstrated that, musicians had more accurate representation of the envelope periodicity of the stimulus than compared to nonmusicians. They noted a high precision of the neural phase locking after long term musical experience is evidenced. The authors also found a correlation between the neural phase locking abilities to the envelope periodicity and years of musical experience. This suggests the role of long term musical experience in shaping the subcortical system which enhances the strong correlation between the duration of the musical training and the neural representation of the features of the stimulus (Lee, Skoe, Kraus & Ashley, 2009).

Pitch coding mechanism in musicians

Neuroimaging studies give important information about how pitch is coded and processed in the human brain. Recording of Frequency Following Response (FFR) also gives information about the neural temporal coding in the brainstem. FFR shows the precision of the temporal pitch information which is dependent on the linguistic and musical experience. With short term training also coding of linguistic pitch can be modified.

The initial temporal pitch coding occurs at the level of auditory periphery and is converted to a code based on the neural firing rate in the brainstem. In the auditory cortex, the information from the individual harmonics of complex tones is combined to form a representation of pitch (Plack, Barker and Hall, 2013).

Frequency Following Response shows the phase locked property which plays a major role in the neural encoding of spectrum and voice pitch of speech sounds. To study this Krishnan, Xu, Grandour and Cariani in the year 2004 recorded the FFR for the set of Chinese monosyllables with rising and falling contours. The FFR was recorded in 13 adult native speakers of Mandarin. The stimulus was presented at 60dBnHL monoaurally at the rate of 3.13/s. The representation of the voice pitch was shown as a phase locked FFR activity. The pitch relevant activity reflected in the FFR is based on the temporal discharge patterns of neurons in rostral brainstem pathways. FFR amplitude for falling tonal sweeps were smaller than the rising tonal sweeps. The FFR representation of harmonics in the dominant region for pitch for the stimulus with rising contour compared to stimulus with falling contours.

The authors Tervaniemi, Just, Koelsch, Widmamm, and Schroger (2005) studied the pitch discrimination accuracy using behavioral measures and event related potential in musicians and non musicians. The 13 professional instrumental musicians who played piano, wind and string instruments and 13 non- musicians were included in the study. For behavioral measure frequency difference threshold was determined and electrophysiological event related potential mismatch negativity (MMN) was administered with standard frequency of 528Hz and deviant frequencies of 532Hz, 539Hz and 550Hz. The results suggest that the MMN amplitude was larger for musicians and also pitch detection was more accurate with the 2% frequency changes and latency was faster in musicians than non-musicians.

Micheyl, Delhommeau, Perrrot and Oxenham (2006) studied the influence of musical training on the pitch discrimination task. The authors considered 68 subjects who had normal hearing sensitivity, among them 18 of them were instrumentalist who played piano, keyboard, flute, violin and viola. Pitch discrimination was found out using both pure tones and complex tones as stimuli and found out the F0 discrimination threshold using two alternative forced choiced procedure. The results suggested that discrimination threshold were 6 times smaller in musicians than those of non- musicians. The F0 difference limen advantage was more seen in complex tones than with pure tones. Because music sounds are harmonic complexes rather than pure tones. Hence, there is enhanced F0 discrimination for complex sounds. The left ear scores were showing an advantage in musicians only. Also, the individuals who played keyboard instrument had large thresholds than those who played other string instruments. Because the individuals who played string instruments had to correctly tune their instruments hence it promotes more accurate pitch discrimination abilities.

Nikjeh, (2006) studied the pitch discrimination using electrophysiological and psychoacoustic procedure in instrumental and vocal musicians. 61 females participated in the study including 19 formally trained vocal musicians, 21 female trained instrumental musicians. Harmonic complexes were used for the pitch discrimination task. The musicians had 6 times more accurate pitch discrimination thresholds than non musicians. There was no significant difference in the pitch discrimination threshold across vocal and instrumental musicians. That is the vocalist and instrumentalist had comparable DLF i.e, 1.4% and 1.3% respectively. Also electrophysiological P1-N1-P2 components were recorded in all the subjects. There was no significant difference between the vocalist and instrumentalist indicating the stimuli is encoded at the level of auditory cortex for group of musicians. The neural responses for musicians occurred earlier than the control subjects. As the MMN is used as a neurological index of discrimination abilities, musicians had earlier latency and larger amplitude than non-musicians for detection of small pitch deviances.

Lee, Skoe, Kraus and Ashley (2009) studied the brainstem responses of two musical interval, the major sixth (E3 and G2) and the minor seventh (E3 and F#2). The participants selected were 26 adults including both musicians (pianists, violinists and vocalists) who had more than 10 years of musical training and non-musicians with less than 3 years of musical training. The results showed that for the consonant and dissonant interval musicians showed larger amplitude for the harmonics of the upper tone E. The amplitude of the harmonics correlated with the numbers of years of experience of musical training. This suggests that musicians showed more accurate representation of the envelope periodicity of the stimulus than non-musicians. The advantage of this study is that it considered the experimental group with more than 10 years of musical training so that the results correlated well with the number of years of experience.

Music and speech are very cognitively demanding auditory phenomena attributed to cortical rather than subcortical circuitry. The stimulus to response correlation ie (Fo contour of stimulus and subject response contours) indicates the faithful procedure to track the pitch and also peak autocorrelation averaged over the entire response gives the useful information about the robustness of the neural phase locking ability with reference to the stimulus. A study done by Wong, Skoe, Russo, Dees and Kraus (2007) to see the brainstem encoding of the linguistic pitch patterns. FFR was recorded in 10 instrumentalists and in 10 non musicians. Mandarin stimuli /mi/ was used as a stimulus which differed only in the Fo. The results showed that the musicians had stronger overall Fo amplitude and FFR amplitude than non musicians. The amplitude of the FFR waveform correlated between musical experience and pitch tracking. This study also provides an evidence for the positive effect of long term music exposure on pitch encoding at the brainstem.

A study done by Musacchia, Sams, Skoe and Kraus in 2007 studied the coding of the pitch and the its harmonics in 29 adults classified further as musicians who played the instrument for more than 10 years and non-musicians without any musical training. The speech evoked auditory brainstem response was recorded for both speech stimuli and musical stimuli. Musicians performed better than non-musicians. Speech evoked fundamental frequency amplitude was consistently correlated with the musical experience in playing the instrument. The fundamental frequency amplitude was also more for the musicians than that of non-musicians group. The authors also found a correlation between the amount of musical practice and strength of the fundamental representation of the pitch. Hence, accurate pitch encoding is important in identifying speakers' message and emotional content of the message. They also observed the top-down influences, originating from complex, multisensory training, and guide plasticity in peripheral areas. This shows that the learning modifies the neural circuitry that governs the performance, beginning with the highest level and gradually refining lower sensory areas.

Pitch was particularly important form of regularity that the auditory system employs to promote auditory object formation and speaker identification (Baumamm & Belin, 2010). Study conducted by Parbey-Clark, Strait and Kraus (2011) in 31 young adults with the age range of 18-30 years, and divided the subjects into 16 musicians who played a musical instrument and 15 subjects as non-musicians. Speech ABR and HINT was carried out. Results obtained indicate that musicians had better perception of speech in noise condition than non-musicians. Musicians also had greater subcortical enhancement of F0 than non-musicians. The degree of enhancement in the auditory brainstem responses to the F0 correlated with Speech in noise perception in both quiet and in noise condition. A positive correlation was found between years of musical experience and extent of F0 enhancement within musicians. The extent of subcortical enhancement to predictably occurring speech in noise relates to ability to hear speech in background noise.

Alexandra, Parbery-Clark, Anderson and Kraus (2013) studied the encoding of fundamental frequency and its harmonics in musicians and nonmusicians. Results showed that musicians had more precise neural encoding of speech in quiet and background noise than nonmusicians. They noted a smaller timing delay in the addition of background noise in musicians. The coding of F0 was greater both in quiet and noise condition in musicians. The addition of noise reduces the neural representation of F0. In the coding of harmonics, musicians did not have greater harmonic encoding than nonmusicians in both quiet and noise condition. The noise also reduces the harmonic encoding.

In the same study, they noted the correlation between the behavioural speech perception scores and the encoding of F0. They found greater encoding of F0 corresponding to better hearing in background noise. The relationship was stronger in the quiet condition than in the noise. Neither the peak latency nor the harmonics did not correlate with behavioural speech perception scores. The conclusion drawn from the study was precise encoding of stimulus in quiet and noise condition suggests more stable representation of the stimulus elements in musicians. The higher levels of response consistency and neural precision in musicians signify the more synchronous neural phase locking ability and hence greater accuracy in neural representation of the speech sound. They also report of the top down process which was promoting the plasticity in the auditory brainstem. This indicates, cortical activity modulates the brainstem activity through the cortico fugal system and that the cortico fugal pathway is important for the auditory learning (Suga, 2008). Musicians also had enhanced top-down control of the subcortical response properties resulting in greater encoding of most behaviourally relevant features which is important for speech perception.

Kraus, Slater, Thompsen, Hornickel, Strait, Nicol and White-Schwoch in 2014 reported a faster latency for V, E and F. The stronger difference in peak E and F correspond to the neural encoding of the rapidly changing consonant transition period After a several months of musical training, faster responses and robust encoding of formant features for the speech signal /da/.

To study the spectral encoding using speech evoked auditory brainstem response in musicians and nonmusicians, Krizman, Slater, Skoe, Marian and Kraus (2015) did a study. They found that F0 encoding was greater in musicians than nonmusicians and harmonics encoding did not differ between musicians and nonmusicians. This suggests the role of F0 in pitch perception is to track an auditory object, attending to a target talker in noise make the F0 as an essential cue for all listeners to attend during the conversation.

Musicians had greater representation of harmonics, earlier peak latency and increased response consistency compared to the nonmusicians (Kraus and Anderson, 2015). The children who had poor performance in the tests of hearing in noise had reduced representation of fundamental frequency.

Kraus, Hornickel, Strait, Slater and Thompson (2014) reported that children who practiced musical instruments more times, had a stronger encoding of speech harmonics. The greater engagement in musical classes predicts stronger speech encoding, which is important for reading and perception.

Children with musical training demonstrated enhanced pitch processing for both music and speech. This highlights the role of musical training in musicians enhances processing of pitch.

Chapter 3

Method

The present study focused on assessing the neural encoding of speech at the brainstem and also to correlate the electrophysiological encoding of fundamental frequency to the behavioral perception of speech.

3.1. Participants

A total of 30 subjects aged between 5-12 years were included”in the study. Participants were divided into two groups. The first group, called as the “instrumental musicians” group, consisted of 15 children who enrolled for instrumental musical training. The second group, called the “non musicians” group, who had 15 children without the instrumental musical training. The musical experience of the instrumental musicians group varied from 6 months to 4 years.

3.1.1. *Inclusion Criteria*

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

- Normal air conduction and bone conduction thresholds (≤15 dB HL) at all octave frequencies from 250 Hz to 8000 Hz
- Normal middle ear function with ‘A’ type tympanogram at 226Hz probe tone with the presence of acoustic reflexes in both ears.
- Speech Recognition Threshold should be within ± 12 dB with respect to the pure tone average of 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz.
- Speech Identification Scores of > 90% at 40 dB SL with respect to the Speech recognition scores.

- Presence of click evoked auditory brainstem response
- No history of neurological or Otological problem
- No reports of illness before or during the testing
- All the participants were native speakers of Kannada

3.2. Test environment

All the tests were carried out in well illuminated, air conditioned rooms which were acoustically treated. The noise levels in the test rooms were within the permissible levels as recommended by ANSI S-3.1 (1991).

3.3. Instrumentation

The following instruments were used in the present study:

- A 2-channel Inventis Piano diagnostic audiometer was used to estimate the hearing thresholds for all the participants and also to assess the speech perception abilities, both in quiet and in noise.
- A calibrated middle ear analyzer GSI-Tympstar was used for tympanometry and reflexometry.
- A Dell laptop was used to deliver the stimulus for SPIN, which were routed through audiometer.
- Auditory brainstem responses were recorded using a Biologic Navigator Pro evoked potential system.

3.4. Stimuli

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

3.5. Procedure

3.5.1. *Pure tone Audiometry*

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

3.5.2. *Speech Audiometry*

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

Speech Identification Scores were measured using the phonetically balanced wordlist of kannada, developed by Yathiraj and Vandana in 1998. The testing done at 40dB SL with reference to the Speech recognition threshold. A total of 25 words were told to the child, the number of correct responses were noted and calculated the percentage.

3.5.3. *Immitance Audiometry*

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

3.5.4. *Speech Perception in Noise (SPIN)*

Speech Perception in Noise test was done using the phonemically balanced (PB) Kannada word list (Yathiraj and Vandana, 1998), recorded in the voice of a typical female Kannada speaker. The stimuli were played in a laptop and were routed through the audiometer. The presentation level was 40 dB SL (with reference to SRT) or at most comfortable level maintain a 0dB SNR. 25 PB words were presented for each ear. The subjects' task was to perceive the words presented in the presence of noise and repeat them back. Each word was given a score of 4 %. Number of correctly identified word at was noted down to find the SPIN score.

3.5.5. *Speech evoked auditory brainstem response*

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

3.5.5.1. Test stimulus

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

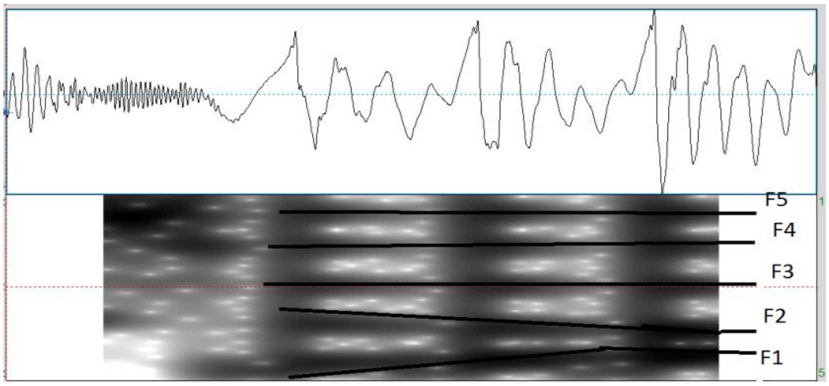

Figure 3:2: Representation of the spectral and temporal aspects of the speech stimulus /da/ used in the present study.

3.5.5.2. Test procedure

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

3.5.5.3. Electrode placement

Initially the electrode sites were cleaned using skin preparation gel (nuPrep). The gold plated disc type electrodes were placed on the scalp at electrode placement site with adequate amount of ten- 20 conduction paste. The electrodes were secured in place using surgical plaster. The testing was done monaurally. The parameters used to record ABR is shown in the table:

The protocol for recording the responses is shown in the tables below:

Table 3.1: Acquisition Parameters for Speech Evoked ABR

Parameters	Target Settings for Speech Evoked ABR in Quiet	Target Settings for Speech Evoked ABR in Noise
Transducer	Insert Earphones	Insert Earphones
Time Window	-10ms to +60 ms	-10 ms to +60 ms
Band-pass filter	70 Hz to 2KHz	70 Hz to 2KHz
No of Channels	One	One
Electrode placement	Non-inverting electrodes(+):Vertex Inverting electrode (–):M1/M2 Ground electrode: Upper Forehead (Fz)	Non-inverting electrodes(+):Vertex Inverting electrode (–):M1/M2 Ground electrode: Upper Forehead (Fz)
Inter-Electrode Impedence	<2 Kilo ohms	<2 Kilo ohms
Sweeps	2000	2000

Table 3.2: Stimulus Parameters for Speech Evoked ABR

Parameters	Target Settings for Speech Evoked ABR in Quiet	Target Settings for Speech Evoked ABR in Noise
Stimulus	/da/	/da/
Masker	None	White Noise
Mode	Ipsilateral	Ipsilateral
Duration of Stimulus	40 ms	40 ms
Polarity	Condensation and rarefaction	Condensation and rarefaction
Stimulus Intensity	80 dBSPL	80 dBSPL
Noise Intensity	N/a	80 dBSPL
Repetition Rate	7.1/second	7.1/second

3.6. Data analysis

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

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

1. The transient response was analyzed in terms of latency and amplitude of V and A peak.
2. The FFR response was analyzed in terms of latency and amplitude of D, E and F peaks (the distance between the peak D, E, F, and O is approximately 10msec which gives the information regarding the encoding of fundamental frequency).

3. The sustained portion was analyzed using Fast Fourier Transformation (FFT) for the latency range of 11.4 ms to 40.6 ms for speech evoked ABR to extract the information regarding the coding of fundamental frequency, first formant frequency and second formant frequency using the MATLAB 2009b version software.

3.6.1. *Procedure for FFT analysis*

To know the coding of fundamental frequency, first formant frequency and higher harmonics, a FFT analysis of the sustained response of the speech evoked ABR was done. This was executed using the MATLAB 2009b version (Brainstem toolbox) developed by Kraus (2004) at Northwestern university. For measuring the fundamental frequency and higher harmonics, Fourier analysis was performed on the 11.4–40.6 ms epoch of the FFR in order to assess the amount of activity occurring over three frequency ranges. Activity occurring in the frequency range of the response corresponding to the fundamental frequency of the speech stimulus (103– 121 Hz), first formant frequencies of the stimulus (454- 719 Hz) and for the higher harmonics (721-1155 Hz) were measured for all the subjects.

Chapter 4

Results

The present study was aimed at understanding the encoding of fundamental frequency in children with instrumental musical training and without musical training. It was also aimed at understanding the correlations between the strength of neural encoding of fundamental frequency and the behavioral perception of speech, both in quiet and in the presence of background noise. For the statistical analyses, the behavioral speech perception scores and the electrophysiological parameters were considered as the dependent variables whereas the different conditions and the sub-conditions i.e., condition (quiet and noise), and ear (predominantly right, predominantly left) as well the two groups (Instrumental musicians and non musicians) were the independent variables.

The brainstem responses were measured by speech evoked auditory brainstem response in quiet and in noise condition. Speech Perception in Noise (SPIN) test was done at 0 dB SNR separately for both the ears. A total of 15 non-musicians and 15 trained Carnatic Instrumental musicians with an age range of 5 to 12 years participated in the study. The data was appropriately tabulated and statistically analyzed using SPSS (version 17) software.

The Following analyses were carried out:

- The mean and standard deviation for the amplitudes of F0, F1 and F2, latency and amplitude of V, A, D, E and F, and SPIN scores in both the groups were calculated using descriptive statistics.
- To find out the test of normality, Shapiro Wilk normality check was done. Few parameters are normally distributed but few were not. The data which was under normal distribution, Multivariate Analysis Of Variance (MANOVA) was carried out for the

group comparison and Independent sample t test for for the comparison of different conditions. The data which were not following normal distribution were subjected to Mann Whitney U test for comparison between groups and Wilcoxon Signed Ranks test for comparison of conditions within the groups.

4.1. Behavioral measures

For behavioral measures, when the SPIN scores for right and left ears were compared between musicians and non-musicians, there was a significant difference seen in the performance between musicians and nonmusicians in both right ear (t(28,15)=4.743, p=0.000) and left ear (/Z/=4.061, p=0.000) respectively.

4.2. Electrophysiological measures:

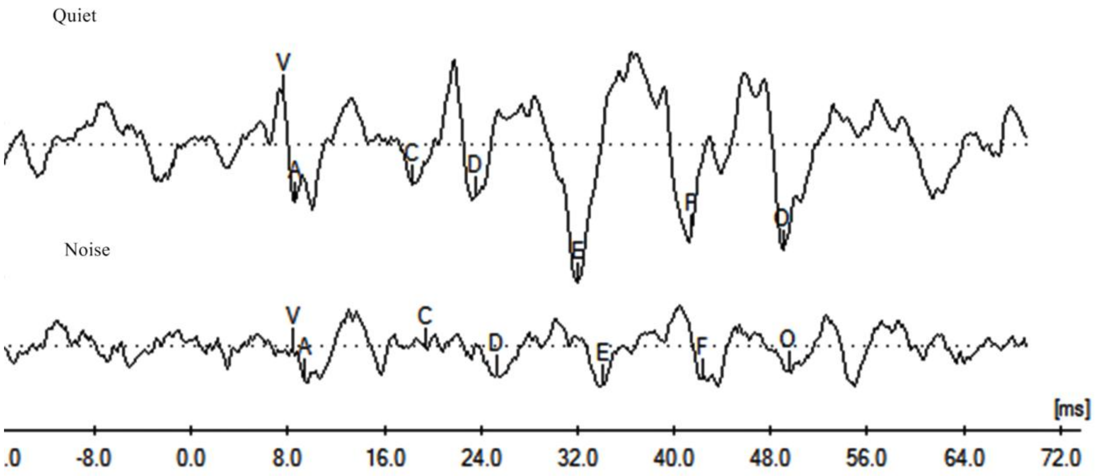

Figure 4.1: Speech evoked auditory brainstem response for quiet and noise condition in musician

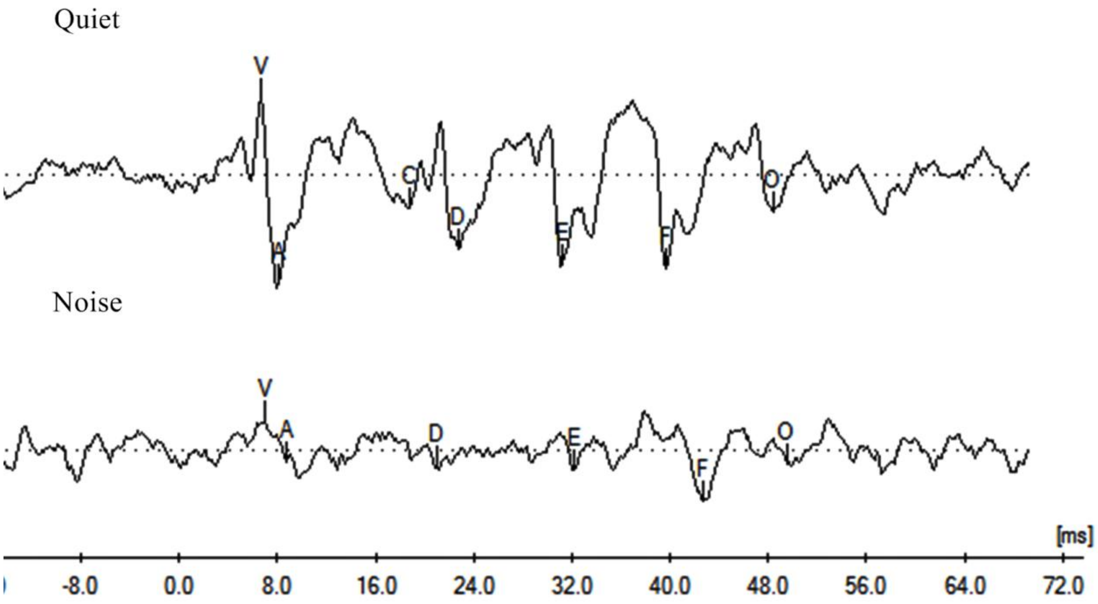

Figure 4.2: Speech evoked auditory brainstem response for quiet and noise condition in nonmusican

4.2.1. *Amplitude of Fundamental frequency*

Statistical analysis was done to compare the frequency of the F0 as encoded at the brainstem level. The frequency was analyzed for each subject across each condition using a MATLAB FFT program.

Table 4.2: Mean and standard deviation for formant amplitudes in micro volts for instrumental musicians and nonmusicians in Quiet condition.

Parameter	Group	Right Ear		Left Ear	
		Mean	SD	Mean	SD
Amplitude of F0	Musicians	10.3463	2.9382	11.6067	2.4543
	Nonmusicians	6.5517	1.8010	8.5828	1.32533
Amplitude of F1	Musicians	1.1960	0.3577	1.2872	0.4224
	Nonmusicians	1.0187	0.4088	1.5045	0.3232
Amplitude of F2	Musicians	0.2886	0.0539	0.4195	0.1907
	Nonmusicians	0.3117	0.12059	0.4233	0.1502

Note. SD= Standard Deviation

Table 4.2: Mean and standard deviation for formant amplitudes in micro volts for instrumental musicians and nonmusicians in Noise condition.

Parameter	Group	Right Ear		Left Ear	
		Mean	SD	Mean	SD
Amplitude of F0	Musicians	4.9913	1.4832	5.5720	1.2408
	Nonmusicians	2.4547	0.6126	2.8675	0.4257
Amplitude of F1	Musicians	0.7173	0.1838	1.1622	1.63756
	Nonmusicians	0.5027	0.2723	0.5085	0.1955
Amplitude of F2	Musicians	0.2213	0.4427	0.2506	0.0662
	Nonmusicians	0.2087	0.0748	0.2218	0.0731

Note. SD= Standard Deviation

In the table 4.1 and 4.2, the mean scores of amplitude of F0, F1 and F2, the musicians had enhanced amplitude of F0, F1 and F2 in both quiet and noise condition than the nonmusicians.

While the amplitudes of F0, F1 and F2 between musicians and non musicians in both quiet and noise condition were compared, it was seen that in quiet condition there was a significant difference between the F0 of both right (F(1,15)=37.477, p=0.000) and left ear (F(1,15)=17.663, p=0.000). The amplitudes of F1 and F2 did not show any statistical significance where p>0.05.

In the noise condition, both amplitude of F0 and F1 of both right and left ears reached statistical significance. The amplitude of F2 does not show any statistical difference between musicians and non musicians.

Table 4.3: Table Statistical test results for comparison of amplitude of F0, F1 and F2 between musician and nonmusicians.

		Test statistic	Sig.
Amplitude of F0	Right ear	6.122	0.000
	Left ear	4.170	0.000
Amplitude of F1	Right ear	6.403	0.017
	Left ear	2.512	0.012
Amplitude of F2	Right ear	0.561	0.579
	Left ear	1.270	0.204

Note: Significance (2 tailed), p< 0.05

Table 4.4: Wilcoxon signed rank test results for amplitude of F0, F1 and F2 of speech ABR for musicians and nonmusicians to compare right and left ear in quiet condition.

	Parameter	Musicians		Nonmusicians	
Quiet condition	Amplitude	/Z/	Significance	/Z/	Significance
Right Vs Left	F0	1.250	0.211	3.237	0.001
	F1	0.341	0.733	3.068	0.002
	F2	2.472	0.013	2.166	0.030

Note. Significance = 2 tailed

In the table 4.4, there was a significant difference seen between the right Vs left ears in quiet condition in the amplitude of F0, F1 and F2 in non-musicians but in musicians there was a significant difference seen in amplitude of F2 alone.

For the comparison between right and left ear amplitude of F0, F1 and F2 in noisy condition, it was seen that there was no significant difference seen in the amplitude of F0, F1 and F2 in musicians where as in non-musicians significant difference was seen only in the amplitude of F0.

Table 4.5: Wilcoxon signed rank test results for amplitude of F0, F1 and F2 of speech ABR for musicians and nonmusicians (quiet and noise condition) for right ear.

	Parameter	Musicians		Nonmusicians	
Right Ear	Amplitude	/Z/	Significance	/Z/	Significance
Quiet Vs Noise	F0	3.408	0.001	3.408	0.001
	F1	3.408	0.001	3.234	0.001
	F2	3.417	0.001	2.858	0.004

Note. Significance = 2 tailed

The results in the table 4.5 showed that there was a significant difference seen in the amplitude of F0, F1 and F2 between quiet and noise condition of right ear in both musicians and nonmusicians.

Table 4.6: Wilcoxon signed rank test results for amplitude of F0, F1 and F2 of speech ABR for musicians and nonmusicians (quiet and noise condition) for left ear.

	Parameter	Musicians		Nonmusicians	
Left Ear	Amplitude	/Z/	Significance	/Z/	Significance
Quiet Vs Noise	F0	3.408	0.001	3.408	0.001
	F1	2.556	0.011	3.409	0.001
	F2	3.170	0.002	3.210	0.001

Note. Significance = 2 tailed

The results in table 4.6 showed that there was a significant difference seen in the amplitude of F0, F1 and F2 between quiet and noise condition of left ear in both musicians and nonmusicians.

4.2.2 *Correlation between behavioral and electrophysiological measures:*

The other aim of the present study was to establish a correlation between the behavioral perception of speech and the neural encoding of speech stimulus at the brainstem level. The data obtained in the behavioral and electrophysiological measure were analyzed using Pearson's correlation coefficient to observe the degree of correlation between behavioral and electrophysiological measure. The behavioral speech perception in noise scores were compared with the amplitude of the F0, F1 and F2.

In both musicians and nonmusicians, SPIN scores did not correlate with the amplitude of F0, F1 and F2 in quiet condition of both right and left ear ($p > 0.05$).

In noise condition, in musicians there was no correlation found in the SPIN scores and amplitude of F0, F1 and F2 ($p > 0.05$). In nonmusicians, the right ear SPIN scores positively correlated with the amplitude of F2 (Correlation Value= 0.621, $p=0.014$) and left ear SPIN scores negatively correlated with the amplitude of F1 (Correlation value= - 0.608. $p=0.016$). This suggests that behavioural SPIN scores reduces when the amplitude of the F1 increases, and SPIN scores increases, when the amplitude of F2 increases.

4.2.2. Comparison of peak latencies

In the table 4.7 and 4.8, the mean scores of the latencies of the peaks ofspeech auditory brainstem was given. In musicians, the latency was earlier compared to nonmusicians.

Table 4.7: Mean and standard deviation for speech evoked auditory brainstem peak latencies in instrumental musicians and nonmusicians in Quiet condition.

Parameter	Group	Right Ear		Left Ear	
		Mean	SD	Mean	SD
Latency of wave V	Musicians	6.968	0.4238	7.1347	0.3071
	Nonmusicians	7.1933	0.4471	7.1493	0.2165
Latency of Wave A	Musicians	8.2453	0.50885	8.1667	0.3627
	Nonmusicians	8.4240	0.5583	8.2073	0.4715
Latency of wave D	Musicians	23.439	1.2216	23.5113	1.0805
	Nonmusicians	24.7860	1.5660	23.4540	0.7975
Latency of wave E	Musicians	32.525	1.2538	32.0300	0.83081
	Nonmusicians	34.1727	1.9526	32.6380	1.70331
Latency of wave F	Musicians	42.1433	2.3576	41.2993	1.3350
	Nonmusicians	44.1400	2.6166	41.8073	2.18408

Note. SD= Standard Deviation

Table 4.8: Mean and standard deviation for speech evoked auditory brainstem peak latencies in instrumental musicians and nonmusicians in Noise condition.

Parameter	Group	Right Ear		Left Ear	
		Mean	SD	Mean	SD
Latency of wave V	Musicians	7.7527	0.7153	7.9080	0.7141
	Nonmusicians	8.2327	0.6087	8.5807	0.9668
Latency of Wave A	Musicians	9.1253	0.8643	9.2173	0.9524
	Nonmusicians	9.5307	0.7510	10.2021	0.6247
Latency of wave D	Musicians	24.8633	1.8604	24.7620	1.7214
	Nonmusicians	25.4573	1.9938	25.1260	1.8817
Latency of wave E	Musicians	34.4713	1.9022	34.2847	0.8882
	Nonmusicians	35.1793	3.0539	33.9193	1.5129
Latency of wave F	Musicians	43.3860	2.30726	43.5193	1.65471
	Nonmusicians	46.1440	2.7265	43.4200	1.37526

Note. SD= Standard Deviation

MANOVA was carried out to see the differences between musicians and nonmusicians for Latencies of speech evoked ABR of right ear. In quiet condition, there was a significant difference seen in latency of E (F (1, 15) = 7.559, p=0.010) and latency of F (F (1, 15) = 4.821, p=0.037). The remaining latencies did not show any statistical difference, where as in the noisy condition, significant difference was seen only in the latency of F (F (1, 15) =0.418, p=0.006). The other parameters did not show any statistical significance where p>0.05 was noted.

While the latencies of Speech ABR of left ear in quiet and in noise condition for musicians and non musicians was compared, in quiet condition, there was no significant difference between musicians and non musicians in the latency where the p was >0.05. In the Noise condition, significant difference was only seen in the latency of A (/Z/= 3.063, p=0.002). The other parameters did not reach statistical significance.

For the group comparison, a significant difference was seen in the noise condition of right ear only (F=2.899, p=0.030).

Table 4.9: Wilcoxon signed rank test results for latency of waves of speech ABR for musicians and nonmusicians (quiet Vs noise condition) for right ear.

	Parameter	Musicians		Nonmusicians	
Right Ear	Latency	/Z/	Significance	/Z/	Significance
Quiet Vs Noise	V	2.983	0.003	3.352	0.001
	A	2.726	0.006	3.125	0.024
	D	2.919	0.004	0.966	0.334
	E	2.869	0.004	1.643	0.100
	F	1.876	0.003	2.104	0.035

Note. Significance = 2 tailed

Table 4.10: Wilcoxon signed rank test results for latency of waves of speech ABR for musicians and nonmusicians (quiet Vs noise condition) for left ear.

	Parameter	Musicians		Nonmusicians	
Left Ear	Latency	/Z/	Significance	/Z/	Significance
Quiet Vs Noise	V	3.154	0.02	3.411	0.001
	A	3.011	0.03	3.238	0.001
	D	1.563	0.118	3.129	0.002
	E	3.298	0.001	2.828	0.005
	F	3.297	0.001	2.484	0.013

Note. Significance = 2 tailed

Wilcoxon signed rank test results reveals that for latency of waves of speech ABR for musicians and non-musicians, there was no significant difference between latencies of right and left ear in musicians. But there was a significant difference seen in the latencies of non-musicians for D (/Z/=3.298, p=0.001), E (/Z/=2.544, p=0.011) and F (/Z/=2.103, p=0.035).

4.2.3. *Comparison of peak amplitudes*

Table 4.11: Mean and standard deviation for speech evoked auditory brainstem peak amplitudes in microvolt for instrumental musicians and nonmusicians in Quiet condition.

Parameter	Group	Right Ear		Left Ear	
		Mean	SD	Mean	SD
Amplitude of wave V	Musicians	0.0980	0.0414	0.1493	0.0516
	Nonmusicians	0.0967	0.0664	0.1493	0.0660
Amplitude of Wave A	Musicians	0.1353	0.08314	0.2367	0.1120
	Nonmusicians	0.0880	0.03342	0.2440	0.05068
Amplitude of wave D	Musicians	0.1247	0.0565	0.2447	0.09320
	Nonmusicians	0.0900	0.03443	0.2680	0.08728
Amplitude of wave E	Musicians	01300	0.0566	0.2680	0.09660
	Nonmusicians	0.0887	0.0253	0.2980	0.07702
Amplitude of wave F	Musicians	0.1427	0.08293	0.2180	0.09526
	Nonmusicians	0.1053	0.03314	0.2420	0.1707

Note. SD= Standard Deviation

Table 4.12 Mean and standard deviation for speech evoked auditory brainstem peak amplitude in microvolt for instrumental musicians and nonmusicians in Noise condition.

Parameter	Group	Right Ear		Left Ear	
		Mean	SD	Mean	SD
Amplitude of wave V	Musicians	0.0604	0.0407	0.0820	0.0270
	Nonmusicians	0.0400	0.0200	0.0553	0.0372
Amplitude of Wave A	Musicians	0.0780	0.0355	0.0693	0.0270
	Nonmusicians	0.0421	0.0269	0.0413	0.0250
Amplitude of wave D	Musicians	0.0840	0.0429	0.1067	0.0900
	Nonmusicians	0.0413	0.0145	0.0500	0.0235
Amplitude of wave E	Musicians	0.0853	0.0348	0.1173	0.04131
	Nonmusicians	0.0380	0.0147	0.0453	0.02031
Amplitude of wave F	Musicians	0.0960	0.0451	0.1287	0.05986
	Nonmusicians	0.0480	0.0214	0.0613	0.0226

Note. SD= Standard Deviation

To compare the amplitude of the speech evoked auditory brainstem response for musicians and nonmusicians, both MANOVA and Mann Whitney U test was carried out. For right ear in quiet condition, there was a significant difference seen in the amplitude of D (F (1, 15) = 4.541, p=0.042) and amplitude of E (/Z/= 2.419, p=0.016). There was no significant difference between musicians and non musicians in the other parameters of amplitude. In the noise condition, the amplitude of A (/Z/=2.692, p=0.007), D (/Z/=3.524, p=0.000) and E (/Z/=3.644, p=0.000) showed a statistical difference between musicians and non musicians. The other peaks of amplitude did not show any statistical significance.

For the left ear comparison between musicians and nonmusicians, in quiet condition, all the amplitude of the peaks did not show any statistical difference (p> 0.05), where as in noisy condition, there was a significant difference seen in the amplitude of A (F(1,15)=10.927, p=0.003), D (F(1,15)=16.886, p=0.000), E (F(1,15)=36.696, p=0.000) and F (F(1.15)=16.602, p=0.000).

Table 4.13: Wilcoxon signed rank test results for amplitude of waves of speech ABR for musicians and nonmusicians to compare quiet Vs noise condition of right ear

Right ear	Parameter	Musicians		Nonmusicians	
	Amplitude	/Z/	Significance (2 tailed)	/Z/	Significance (2 tailed)
Quiet Vs Noise	V	2.415	0.16	3.117	0.002
	A	3.178	0.001	3.072	0.002
	D	2.225	0.026	3.419	0.001
	E	2.520	0.012	3.417	0.001
	F	1.845	0.065	3.340	0.001

In the table 4.14, there is a significant difference seen in the amplitudes of the non-musicians for quiet Vs noise condition in right ear, where as in musicians amplitude of wave V and F did not show any significant difference between quiet and noise condition.

Table 4.14: Wilcoxon signed rank test results for amplitude of waves of speech ABR for musicians and nonmusicians (quiet Vs noise condition) of left ear

Left ear	Parameter	Musicians		Nonmusicians	
	Amplitude	/Z/	Significance	/Z/	Significance
Quiet Vs Noise	V	3.326	0.001	3.239	0.001
	A	3.297	0.001	3.413	0.001
	D	3.299	0.001	3.409	0.001
	E	3.235	0.001	3.411	0.001
	F	2.973	0.003	3.409	0.001

Note: significance- 2 tailed

In the table4.15, there was a significant difference noted in both musicians and nonmusicians for quiet Vs noise condition in left ear for amplitude of Speech ABR peaks.

Table 4.15: Wilcoxon signed rank test results for amplitude of waves of speech ABR for musicians and nonmusicians (right and left ear) in quiet condition

Quiet	Parameter	Musicians		Nonmusicians	
	Amplitude	/Z/	Sig.(2 tailed)	/Z/	Sig.(2 tailed)
Right Vs Left	V	2.487	0.013	2.908	0.004
	A	2.615	0.009	3.414	0.001
	D	3.068	0.002	3.412	0.001
	E	2.842	0.004	3.409	0.001
	F	2.042	0.041	3.298	0.001

There was a significant difference seen in the amplitude of Speech ABR for right ear Vs left ear in quiet condition for the group’s musicians and non-musicians.

When the right ear Vs left ear scores in the noise condition was compared, it was observed that in musicians there was a significant difference seen only in the amplitude of V (/Z/=2.267, p=0.023) and E (/Z/=2.047, p=0.041). In non-musicians there was significant difference seen only in the amplitude of A (/Z/=0.315, p=0.001). The remaining amplitudes of the Speech evoked auditory brainstem response did not show any significant difference.

Chapter 5

Discussion

The aim of the present study was to see the neural encoding of the pitch in children who are learning Carnatic instrumental music. The objectives of the study were to investigate the pitch encoding mechanism in children with instrumental Carnatic musical training using speech evoked auditory brainstem responses and to observe the relation between pitch encoding mechanism and speech perception abilities in musicians. The results obtained from different statistical analyses for the groups are discussed below:

5.1. Comparison of the amplitude of the fundamental frequency and the harmonics

In the present study the comparison of the amplitude of F0, F1 and F2 between musicians and nonmusicians, showed a significant difference seen in the Fundamental frequency (F0) of both right ear and left ear in quiet condition. The amplitude of F1 and F2 did not show any significant difference. These findings are supported by the results of Wong, Skoe, Russo, Dess and Kraus (2007), they reported musicians had stronger overall F0 amplitude than the nonmusicians. It is also supported by Musacchia, Sams, Skoe and Kraus (2007), they reported, the amplitude of the F0 was larger in musicians than the control group. This result is also correlated with the musical experience in playing the instrument (Lee, Ske, Kraus and Ashley,2009). This shows the more accurate representation of the envelope periodicity of the stimulus seen in the musicians than the nonmusicians. This could be due to the neural precision and more synchronous neural phase locking ability seen in musicians. Due to the plasticity seen in musicians due to the musical training, enhanced top-down processing is evidenced which results in encoding of the stimulus related features such as F0, F1 and F2 (Alexandra, Parbery-Clark, Anderson and Kraus, 2013).

Whereas in noise condition, there was a significant difference seen in the amplitude of F0 and F1 in both right and left ears. The amplitude of F2 did not show any significant difference between musicians and nonmusicians. This finding is supported by the study carried out by Parbery-Clark, Anderson and Kraus (2013), they found that the musicians had precise neural encoding of the speech in both quiet and noise conditions. This indicates a greater encoding of f0 in both quiet and noise conditions. The musicians did not have greater harmonic encoding than nonmusicians. In quiet condition, the this is replicated in the present study also.

In comparison of conditions, quiet and noise in both musicians and nonmusicians, there was a significant difference seen in the amplitude of F0, F1 and F2 in both musicians and nonmusicians. These results are supported by Russo, Nicol, Musacchia and Kraus (2004), they reported that the amplitude of F0, F1 and F2 are also affected by the presence of background noise. This shows a robust encoding of fundamental frequency and its formants in the quiet condition than the noise condition. This can be attributed to the structural changes occur in the auditory brainstem of the children who begin musical training in early life, these can be genetically predisposed to have the robust auditory brainstem functions (Strait, Parbery-Clark, Hittner and Kraus, 2012).

In quiet condition, there was no significant correlation found between the behavioral SPIN scores and the amplitude of F0, F1 and F2 for both musicians and nonmusicians. The result of the present study was supported by Alexandra, Parbery-Clark, Anderson and Kraus (2013). They also found no correlation between the behavioral speech in noise scores and the amplitude of F0, F1 and F2. Where as in noise condition, SPIN scores did not correlated with the amplitude of F0, F1 and F2 in musicians. In nonmusicians, SPIN scores positively correlated with the amplitude of F2 (right ear) and negatively correlated with the amplitude of F1. Where as in contrast to the result of the present study, Musacchia, Sams, Skoe and Kraus (2007), noted a correlation between the behavioral speech perception scores and amplitude of F0. This could be due to the years of musical experience, which strengthen the representation of the pitch in the auditory brainstem.

5.2.Comparison of the peak latency

In the present study, the latency of the speech evoked auditory brainstem response, V, A, D, E and F were compared between the instrumental musicians and nonmusicians. The results showed that in quiet condition, there was no significant difference seen in the latency of left ear, whereas in right ear, significant difference was seen in the latency of E and F. These results can be supported by the study of Parbery-Clark, Skoe and Kraus (2009) for the results of left ear. They reported that, the latency of the speech evoked auditory brainstem response recorded in quiet condition in musicians and nonmusicians can have equivalent peak latency. But in contrast to above study, the peak latencies of the right ear showed significant delay in the latency of the E and F in nonmusicians compared to musicians. This is also supported by the other study done by the Bidelman and Alain (2015), reported that musicians had earlier peak latency than their nonmusicians peers. This is due to the musical training, which shows faster and robust neural encoding at the level of brainstem and cortex.

In noise condition, the results showed a significant difference seen only in the latency of F in right ear and latency of A in left ear. This is in agreement with the study done by the authors Parbery-Clark, Skoe and Kraus (2009) which reported that, musicians are less sensitive to the degradative effects of noise, and the difference is seen at the onset of the peak latencies. Russo, Nicol, Musacchia and Kraus (2004) report that, in the presence of noise, the transient peaks were less degraded in the presence of noise than the onset responses. Latency V and A were most affected by the presence of background noise than the latency of D, E and F. This is also supported by the study done by White-Schwoch, Carr, Anderson, Strait and Kraus (2013). However, in the present study also, the latency of the onset response A showed a significant delay in the nonmusicians than compared to instrumental musicians.

In comparison of quiet Vs noise conditions in the groups, a significant difference was seen in the latencies of the speech evoked auditory brainstem response V, A, D, E and F in both quiet and noise condition for musicians. In nonmusicians, there was significant difference seen in the latencies of V, A and F of right ear and V, A, D , E and F of the left ear. The findings of the present study are in agreement with the study of Parbery-Clark, Skoe and Kraus (2009) and Russo, Nicol, Musacchia and Kraus (2004). They reported that, in the presence of noise there is a delay in the latency of speech evoked auditory brainstem response. The result also supported by the study done by White-Schwoch, Carr, Anderson, Strait and Kraus (2013), this indicates the disruptive effects of noise on the latency of the speech evoked auditory brainstem response. This could be due to the enhancement of strengthened top-down auditory mechanism.

5.3. Comparison of peak amplitude

The amplitude of speech evoked Auditory brainstem response were compared between musicians and nonmusicians in the present study. In quiet condition, there was no significant difference seen in the amplitude between musicians and nonmusicians in left ear and significant difference was seen in the amplitude of D and E in the right ear. This is supported by the study conducted by Parbery-Clark, Skoe and Kraus (2009). These authors reported that there was no significant difference seen in the amplitude of the speech evoked auditory brainstem response in quiet condition between musicians and nonmusicans.

In noise condition, significant difference was seen in the amplitude of A, D, E and F of the left ear and amplitude of A and E in the right ear between musicians and nonmusicians. This is in agreement with the study done by Zubin and Rajalakshmi (2012). From the studies we know that, the amplitude of the speech evoked auditory brainstem response is variable across the subjects (Parbery-Clark et al, 2009).

In comparison, of quiet vs noise condition, there was a significant difference seen in the peak amplitude of V, A, D, E and F in both musicians and nonmusicians for the left ear. In right ear, significant difference was seen in all the amplitude of speech evoked auditory brainstem response in nonmusicians and only in amplitude of A, D and E in musicians. The above findings are supported by the study done by Russo, Nicol, Musacchia and Kraus (2004).

Chapter 6

Summary and Conclusion

The present study was aimed to study the neural encoding of the pitch in children with instrumental musicians and nonmusicians. A total of 15 children who learnt Carnatic instrumental musicians participated in the study. The coding of pitch and the harmonics were measured using Speech evoked auditory brainstem response. Behavioral speech perception scores was measured at 0 dB SNR. Speech evoked auditory brainstem response was recorded both in quiet and noise condition. Fast Fourier Transform was done to find the energy concentration at the fundamental frequency and its formants.

The following conclusions are drawn from the results of the present study:

1. F0 coding is better in instrumental musicians than the nonmusicians.
2. The coding of the fundamental frequency and the harmonics were better in the quiet condition than the noise condition.
3. The SPIN scores did not correlate with the amplitude of the F0, F1 and F2 in quiet condition.
4. In noise condition, musicians behavioral speech perception scores did not have any correlation with the SPIN scores. In nonmusicians, SPIN scores positively correlated with the amplitude of F2 in right ear and negatively correlated with the amplitude of F1 in left ear.
5. The latencies were noted to be earlier in the instrumental musicians than the nonmusicians.
6. There was a significant difference seen in the latency recorded both in quiet and noise condition.

7. The amplitude were robust in the instrumental musicians than the nonmusicians.
8. Compared to quiet and noise condition, the amplitude is greater in the quiet condition than the noise condition.

Implications:

1. This study has added information to the existing literature.
2. The results of the study indicate the utility of speech ABR for the encoding of pitch in children.
3. Musical training can be implemented early in the childhood to overcome the adverse effects of the noise.

Future Directions for research:

1. Future research can be carried out on between Hindustani and Carnatic musicians.
2. Musicians and dancers can be compared to study the encoding of the F0 and also to find out the ability to perceive the speech in noise.

References

Anderson, S., Skoe, E., Chandrasekaran, B., Zecker.S., & Kraus, N. (2010). Brainstem correlates of speech-in-noise perception in children. *Journal of hearing research* (270), 151-157.

Anderson, S., White-Schwoch, T., Parbery-Clark, A., & Kraus, N. (2013). Reversal of age-related neural timing delays with training. *Proceedings of the National Academy of Sciences*, *110*(11), 4357-4362.

Bajo, V. M., Nodal, F. R., Moore, D. R., & King, A. J. (2010). The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nature neuroscience*, *13*(2), 253-260.

Baumann, O., & Belin, P. (2010). Perceptual scaling of voice identity: common dimensions for different vowels and speakers. *Psychological Research*, *74*(1), 110-120.

Bidelman, G. M., & Alain, C. (2015). Musical training orchestrates coordinated neuroplasticity in auditory brainstem and cortex to counteract age-related declines in categorical vowel perception. *The Journal of Neuroscience*, *35*(3), 1240-1249.

Bidelman, G. M., & Krishnan, A. (2010). Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain research*, *1355*, 112-125.

Bidelman, G.M., Krishnan, A., & Gandour, J.T. (2010). Enhanced brainstem encoding predicts musicians' perceptual advantages with pitch. *European journal of neuroscience* (33), 530-538.

Carhart, R., & Jerger, J. (1959). Preferred method for clinical determination of pure-tone thresholds. *Journal of Speech and Hearing Disorders (24*), 330-345.

Francois, C., & Schon, D. (2011). Musical expertise boosts implicit learning of both musical and linguistic structures. *Cerebral Cortex* (21), 2357-2365.

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

Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., & Schlaug, G. (2009). Musical training shapes structural brain development. *The Journal of Neuroscience*, *29*(10), 3019-3025.

Kraus, N., & Slater, J. (2015). Music and language: relations and disconnections. *Handbook of clinical neurology, 129*- 207.

Kraus, N., & Anderson, S. (2014). Community-Based Training Shows Objective Evidence of Efficacy. *The Hearing Journal*, *67*(11), 46-48.

Kraus, N., & Anderson, S. (2015). Identifying Neural Signatures of Auditory Function. *The Hearing Journal*, *68*(1), 38-40.

Kraus, N., & Anderson, S. (2015). The ear-brain connection: the role of cognition in neural speech processing.

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, *11*(8), 599-605.

Kraus, N., Hornickel, J., Strait, D. L., Slater, J., & Thompson, E. (2014). Engagement in community music classes sparks neuroplasticity and language development in children from disadvantaged backgrounds. *Frontiers in psychology*, *5*.

Kraus, N., Slater, J., Thompson, E. C., Hornickel, J., Strait, D. L., Nicol, T., & White-Schwoch, T. (2014). Auditory learning through active engagement with sound: biological impact of community music lessons in at-risk children.*Frontiers in neuroscience*, *8*.

Krishnan, A., Xu, Y., Gandour, J. T., & Cariani, P. A. (2004). Human frequency-following response: representation of pitch contours in Chinese tones. *Hearing Research*, *189*(1), 1-12.

Krizman, J., Slater, J., Skoe, E., Marian, V., & Kraus, N. (2015). Neural processing of speech in children is influenced by extent of bilingual experience. *Neuroscience letters*, *585*, 48-53.

Lee, K. M., Skoe, E., Kraus, N., & Ashley, R. (2009). Selective subcortical enhancement of musical intervals in musicians. *The Journal of Neuroscience*, *29*(18), 5832-5840.

Klatt, D. H. (1980). Software for a cascade/parallel formant synthesizer. *The journal of the Acoustical Society of America*, 67(3), 971-995.

Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing research*, *219*(1), 36-47.

Moreno, S., & Bidelman, G. M. (2014). Examining neural plasticity and cognitive benefit through the unique lens of musical training. *Hearing research,308*, 84-97.

Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S.L., & Besson, M. (2009). Musical Training Influences Linguistic Abilities in 8-Year-Old Children: More Evidence for Brain Plasticity. *Cerebral Cortex* (19), 712-723.

Musacchia, G., Sams, M., Skoe, E., & Kraus, N. (2007). Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences*, *104*(40), 15894-15898.

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 the United States* (104), *15894-15898.*

Nikjeh, D. A. (2006). Vocal and instrumental musicians: Electrophysiologic and psychoacoustic analysis of pitch discrimination and production.

Ohnishi, T., Matsuda, H., Asada, T., Aruga, M., Hirakata, M., Nishikawa, M., & Imabayashi, E. (2001). Functional anatomy of musical perception in musicians. *Cerebral Cortex, 11*(8), 754-760.

Parbery-Clark, A., Anderson, S., & Kraus, N. (2013). Musicians change their tune: how hearing loss alters the neural code. *Hearing research*, *302*, 121-131.

Parbery-Clark, A., Anderson, S., Hittner, E., & Kraus, N. (2012). Musical experience strengthens the neural representation of sounds important for communication in middle-aged adults. *Frontiers in aging neuroscience, 4*.

Parbery-Clark, A., Anderson, S., Hittner, E., & Kraus, N. (2012). Musical experience offsets age-related delays in neural timing. *Neurobiology of aging, 33*(7), 1483-e1.

Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical experience limits the degradative effects of background noise on the neural processing of sound. *The Journal of Neuroscience*, *29*(45), 14100-14107.

Parbery-Clark, A., Strait, D. L., & Kraus, N. (2011). Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. *Neuropsychologia*, *49*(12), 3338-3345.

Parbery-Clark, A., Tierney, A., Strait, D. L., & Kraus, N. (2012). Musicians have fine-tuned neural distinction of speech syllables. *Neuroscience*, *219*, 111-119.

Plack, C. J., Barker, D., & Hall, D. A. (2014). Pitch coding and pitch processing in the human brain. *Hearing research*, *307*, 53-64.

Rajalakshmi, K., & Thomas, J. (2012). Periodicity Coding and Speech Perception in Noise in Individuals with Symmetrical And Symmetrical cochlear hearing loss. Unpublished Independent Project submitted to AIISH.

Rajalakshmi, K. (2015). *Music & Hearing*, Nova Sciences Publishers, Inc, NewYork.

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

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

Shahin, A. J. (2011). Neurophysiological influence of musical training on speech perception. *Frontiers in psychology*, *2*, 1-10.

Sinha, S. K., & Basavaraj, V. (2010). Speech evoked auditory brainstem responses: a new tool to study brainstem encoding of speech sounds. *Indian Journal of Otolaryngology and Head & Neck Surgery*, *62*(4), 395-399.

Skoe, E., & Kraus, N. (2010). Auditory brainstem response to complex sounds: a tutorial. *Ear and hearing*, *31*(3), 1-23.

Skoe, E., & Kraus, N. (2012). A Little Goes a Long Way: How the Adult Brain Is Shaped by Musical Training in Childhood. *The Journal of Neuroscience*, (34), 11507–11510.

Song, J. , Skoe, E., Banai, K.., & Kraus, N. (2011). Perception of speech in noise: Neural correlates. *Journal of Cognitive Neuroscience,* 23(9), 2268-2279.

Song, J. H., Skoe, E., Banai, K., & Kraus, N. (2011). Training to improve hearing speech in noise: biological mechanisms. *Cerebral Cortex*, *22*(5), 1180-90.

Strait, D. L., & Kraus, N. (2014). Biological impact of auditory expertise across the life span: musicians as a model of auditory learning. *Hearing research*, *308*, 109-121.

Strait, D. L., Chan, K., Ashley, R., & Kraus, N. (2012). Specialization among the specialized: auditory brainstem function is tuned in to timbre. *Cortex*, *48*(3), 360-362.

Strait, D. L., Parbery-Clark, A., O'Connell, S., & Kraus, N. (2013). Biological impact of preschool music classes on processing speech in noise. *Developmental cognitive neuroscience*, *6*, 51-60.

Strait, D. L., Slater, J., O'Connell, S., & Kraus, N. (2015). Music training relates to the development of neural mechanisms of selective auditory attention. *Developmental Cognitive Neuroscience*, *12*, 94-104.

Suga, N. (2008). Role of corticofugal feedback in hearing. *Journal of Comparative Physiology*, *194*(2), 169-183.

Tervaniemi, M., Just, V., Koelsch, S., Widmann, A., & Schröger, E. (2005). Pitch discrimination accuracy in musicians vs nonmusicians: an event-related potential and behavioral study. *Experimental brain research, 161*(1), 1-10

Tierney, A., Krizman, J., Skoe, E., Johnston, K., & Kraus, N. (2013). High school music classes enhance the neural processing of speech. *Frontiers in psychology, 4*.

Vandana, S., & Yathiraj, A. (1998). Speech identification test for kannada speaking children. Unpublished independent project done at AIISH.

White-Schwoch, T., Carr, K. W., Anderson, S., Strait, D. L., & Kraus, N. (2013). Older adults benefit from music training early in life: Biological evidence for long-term training-driven plasticity. *The Journal of Neuroscience*, *33*(45), 17667-17674.

Wong, P. C. M., 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.

Zubin, V., & Rajalakshmi, K. (2012). Brainstem correlates of speech perception in noise: Carnatic Musicians versus Non-Musicians, Published Master’s Dissertation, submitted to University of Mysore, Mysore.