Subcortical Encoding of Manipuri Pitch Contours inNative Speakers and Non-Native Speakers

[1]Sneha Bansal & [2]Vijaya Kumar Narne

Abstract

The aim of the present study was to examine the subcortical representation of pitch contours of Manipuri language in native Manipuri-speakers and in non-native speakers with and without musical training. The study also aimed to compare the FFRs recorded in native Manipuri speakers, with and without musically trained, non-native speakers. Three Manipuri disyllabic words were spoken by a native speaker and the three tonal contours (rising, falling and flat) were extracted. These were used as stimuli to extract the FFRs in the three groups of subjects. Pitch tracking accuracy and pitch strength were calculated of the response obtained from each of the groups. Overall, the results demonstrated that the pitch strength was significantly better in the Manipuri group, followed by musically trained group and then the non-manipuri group for rising and flat tones.

Keywords: *Frequency Following Response, tonal contours, Pitch tacking accuracy, Pitch strength.*

Introduction

Tonal languages are those in which pitch variations are used to indicate different meanings at the word or syllable level.Such languages are common in the Far East, South-East Asia (i.e Thai, Cantonese, Mandarin and Taiwanese) and West Africa (Yoruba). Languages in which pitch variations are usually not leading to different meanings at the syllable or word level (for example, English, Hindi, and Kannada) are called as non-tonal languages. In these languages pitch variations indicate stress, different intonation patterns at post lexical levels. Tonal languages provide an optimal window for investigating how long-term experience with time-varying pitch patterns shape perceptual and neural processing of pitch.

Earlier investigators believed that the language processing and plasticity are confined to the cortical structures but they are not a part of the brainstem encoding (Hickok &Poeppel, 2004; Zatorre, Evans, Meyer, &Gjedde, 1992).Speech specific operations may not begin until the signal reaches the cerebral cortex (Scott &Johnsrude, 2003).

However, Galbraith, Arabagey, Branski, Commerci and Rector(1995), Galbraith et al. (2004), showed increased amplitude of Frequency Following Response (FFR) to forward speech when compared to reversed speech.

These results indicate the familiar phonetic and prosodic properties of forward speech that selectively activates the brainstem neurons. It is hence now being accepted that subcortical structures may contribute actively to auditory processing, and are not simply passive relay stations sending information from the peripheral sensory organs to the cortex. In addition, FFR studies have demonstrated that even at the early subcortical stages of processing, the auditory system is found to be malleable due to interactions between the sensory and cognitive processes (Kraus &Banai, 2007). Thus, over the years, evidence is mounting that the experience dependent neural plasticity is not only limited to the auditory cortex but is also seen in the subcortical structures (Griffiths, Uppenkamp, Johnsrude, Josephs & Patterson, 2001).

Krishnan, Xu, Gandour and Cariani (2005) measured the impact of long-term language experience on pitch coding using frequency following response (FFR). They found that native Mandarin-speaking subjects, with an exposure of about twenty years showed more precise linguistic pitch pattern encoding relative to the native English-speaking subjects. Speakers of tonal (Mandarin) language have improved pitch encoding at the level of brainstem compared to non-tonal (English) language speakers irrespective of whether it is a speech or non-speech context (Krishnan, Swaminathan&Gandour, 2009; Swaminathan, Krishnan &Gandour, 2008).

It is well documented that neural encoding of pitch in the auditory brainstem is shaped by long-term experience with language and music (acoustic experience). FFRs between native Mandarin speakers and English speakers have shown that long-term experience with lexical tones improves periodicity encoding at the brainstem (Krishnan, Swaminathan&Gandour, 2009; Krishnan, Xu, Gandour&Cariani, 2005; Swaminathan, Krishnan &Gandour, 2008). These investigators have proposed that the precise pitch pattern encoding is specific to language experience but not an acoustic experience. However, few other investigators, Wong, Skoe, Russo, Dees and Kraus (2007) found that English (non-tonal language) speaking musicians who did not speak Mandarin (tonal-language) showed more robust and faithful encoding of Mandarin tones. The aforementioned studies indicate that both language and acoustical experience (music) would result in a more robust encoding of

[1]Email: snehsmiles@gmail.com,
[2]Lecturer in Audiology, Email: vijaynarne@gmail.com

signal at the brainstem.

According to the studies, tonal languages and music arguably provide the best window for studying how long-term experience with time-varying pitch patterns shapes the neural representation of periodicity. There are only a handful of studies which have examined the extent to which experience dependent effect is specific to a particular language or a general acoustical experience. Also, there are few studies investigating subcortical processing in the tonal languages of India. Therefore, there is a need to examine the long term effects of language exposure and music exposure in the Indian context.

In the present study, Dynamic Iterated Rippled Noise (DyIRN) has been used as the stimuli which allow us to study the neural mechanisms underlying the representation of pitch patterns of those that occur in natural speech without a semantic confound. In the present study, FFRs of native and non-native (with no exposure to music) speakers of Manipuri were recorded and compared to assess the effect of long-term experience with language. The same was carried out with FFRs recorded from musically trained non-native speakers with native speakers of Manipuri, who have not been trained in music to assess if musical experience confers similar benefits as tonal language exposure.

Method

Participants

In the present study, Frequency Following Response (FFR) was recorded from three group of subjects. These groups were native Manipuri speakers as Group I, musically trained non-Manipuri speakers as Group II, and non-musically trained non-Manipuri speakers. All of them had pure tone thresholds within 15 dB HL at octave frequencies between 250 Hz and 8000 Hz. A normal middle ear function was ensured with type-A tympanogram and the presence of bilateral acoustic reflexes. Also, a detailed case history ensured that none of them had any history of middle ear pathology, and did not have complaint of any neurological problem.

Group-I consisted of 10 individuals (5 males & 5 females), who were native speakers of Manipuri, a tonal language spoken in a north-eastern state of India were included in group I. All the subjects in this group were born and raised in Manipur. All of them were in the age range of 18-25 years with mean age 22 years, 3 months.

Group-II consisted of 10 individuals (5 males & 5 females), who were speakers of a non-tonal language were included in group II. None of the subjects in this group were exposed to any tonal language and had no formal musical training. All of them were in the age range of 18-25 years with mean age 19.2 years.

Group-III consisted of 10 participants (3 males & 7 females), who were speakers of a non-tonal language were included in group III. None of the subjects in this group were exposed to any tonal language. They had undergone formal music training for a minimum of five years. All of them were in the age range of 18-25 years with the mean age of 21.1 years.

Test Stimuli

Three disyllabic words of Manipuri were used for recording FFRs. All of the three words were phonetically similar but were distinguished by their tonal contours having three different intonation patterns, rising (Tc1), flat (Tc2), falling (Tc3), signaling three different meanings. The three tonal variations of the stimuli were /tʃaba/-1, /tʃaba/-2 and /tʃaba/-3 which in Manipuri means, "swimming", "eating" and "fitting" respectively. As the tonal variations are lexical for Manipuri-speakers and not for non-Manipuri-speakers, these words were used as the stimuli to test the objectives of the study.

These three words were spoken by an adult female native Manipuri-speaker. They were recorded using a directional microphone in Praat software (version 5.1.31) at a sampling frequency of 44.1 kHz and 16 bit analog to digital convertor. All the stimuli were recorded in a sound treated room.

The three different pitch patterns were extracted using short-term autocorrelation in MATLAB 7. The pitch contours of the three different tones are represented in Figure 1. A polynomial fitting function was employed to fit pitch contour and the equation generated was used in generation of dynamic iterated ripple noise (DyIRN).

Each of these DyIRNs had duration of 250 ms, which included a 10ms cosine ramp to eliminate the spectral

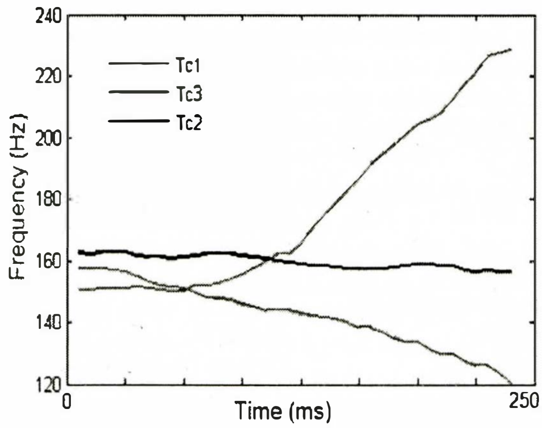

Figure 1: Frequency variations as a function of time of the three different tonal contours, (i.e. Tc1 (Rising), Tc2 (Flat) and Tc3 (Falling)) of the stimulus used in the study.

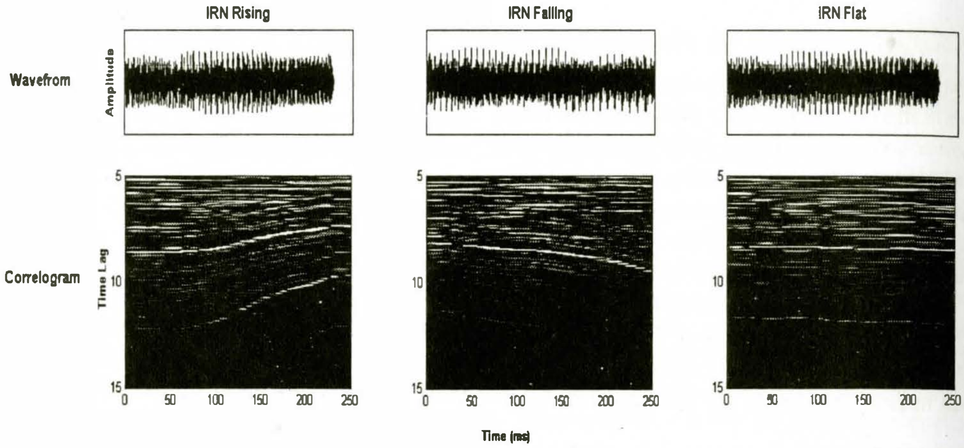

Figure 2: Stimulus waveforms (row 1) and correlograms (row 2), in order from top to bottom, of the IRN homolog of the three different Manipuri tonal contours (rising, falling and flat) of the stimulus used in the study.

splatter and to minimize the onset responses. All the time-varying IRN stimuli were created at a high iteration step (n=12), with a gain of 0.9. Figure 2 shows the waveform, spectrogram and correlogram of the stimuli used in the present study. The generation of IRN noise is provided in the Appendix 1. The stimuli were root mean square normalized to maintain uniform amplitude across all the three stimuli, using group normalization in Adobe audition software (version 3.0). Each stimulus was then converted to a STM file, using Intelligent Hearing System (IHS) stimulus conversion software.

Recording of FFR

Participants were made to relax (preferably sleep) on a reclining chair and they were asked to refrain from any extraneous body movements to minimize movement artifacts. FFRs were recorded from a single channel using gold plated electrodes placed at Cz and referenced to the tip of the nose. An electrode at contralateral mastoid served as ground. The sites of electrode placement were prepared with skin preparation gel and the electrodes were held in their respective positions with a plaster. The absolute electrode impedances were maintained below 5 kΩ and the inter-electrode impedances were less than 2 kΩ in order to maintain low level of noise. The human (FFR) was recorded using the parameters listed in Table 1.

Table 1: Stimulus and acquisition parameters for recording of FFT

Parameters	Stimulus
Duration of stimuli	250 ms
Intensity	70 dBSPL
Transducer	Eartone-3A-insert earphones
Repetition rate	2.76 per second
Number of samples	2000
Filter setting	100 - 3000 Hz
Analysis window	300ms
Mode of presentation	Monoaural
Gain	2,00,000

The order of the stimuli presentation was randomized across the participants. The stimulus files were routed through a digital to analog module and then presented monaurally to the test ear through the magnetically shielded insert earphone (IHS, ER-3A). Each stimulus was presented twice and the stimulus locked responses were subsequently averaged for each stimulus condition to yield a response with a higher signal-to noise-ratio.

Data Analysis

The data obtained was analyzed for the pitch tracking accuracy and the pitch strength of the Manipuri tonal contours.

The ability of the FFR to follow the pitch changes in the stimuli was evaluated by extracting the fo contour from the FFRs using a periodicity detection short term autocorrelation algorithm (Boersma, 1993). Essentially, the algorithm works by sliding a 40 ms window in 10 ms increments over the time course of the FFR. The autocorrelation function will be computed for each 40ms frame and the time- lag corresponding to the maximum autocorrelation value within each frame will be recorded. The reciprocal of this time-lag (or pitch period) represents an estimate of fo. The time- lags associated with autocorrelation peaks for each frame were concatenated together to give a running fo contour. This analysis will be performed on both the FFRs and their corresponding stimuli. Pitch tracking accuracy is computed as the cross- correlation coefficient between the fo contour extracted from the FFRs and the fo contour extracted from the stimuli.

To compute the pitch strength of the FFRs to time varying stimuli, FFRswere divided into six non overlapping 40ms sections (5-45ms; 45-85; 85- 125; 125- 165; 165-

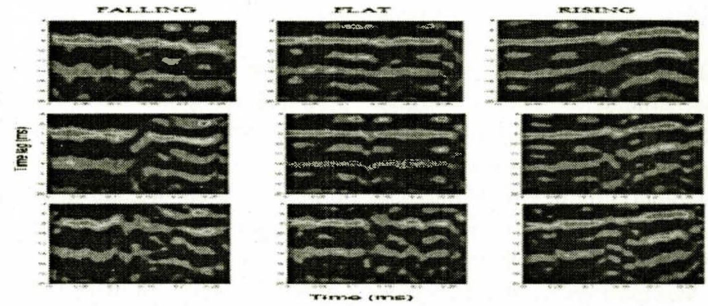

Figure 3: Correlograms derived from the grand average FFR waveforms for the Manipuri group (panel 1), musically trained group (panel 2) and non-tonal group (panel 3) for falling followed by flat and rising tone.

205;205-245ms). The normalized autocorrelation function (expressed as a value between 0 and 1) was be computed for each of these sections, from an analysis of corresponding time frames of the three stimuli and their FFR responses. Here "0" represented an absence of periodicity and "1" represented maximal periodicity. Within each 40ms section, a response peak was selected which would correspond to the same location (time-lag) of the autocorrelation peak in the input stimulus (Krishnan et al., 2009; Krishnan et al., 2009; Swaminathan et al., 2008). The magnitude of this response peak represented an estimate of the pitch strength per section. Pitch strength was measured by the average magnitude of the normalized autocorrelation peak per language in each 40 ms frame. All the data analyses were performed using custom routines coded in MATLAB 7 (The Math Works, Inc., Natick, MA). Appropriate statistical analyses were done using SPSS 20.

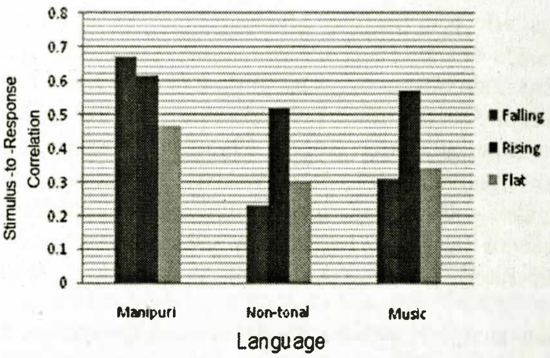

Figure 4: Mean pitch tracking accuracy of FFR for three Manipuri tonal contours (i.e falling, rising and flat) across the three groups of subjects (i.e Manipuri speakers, non-tonal language speakers and musically trained non-tonal speakers).

Results

Temporal and Spectral Properties

Correlograms for grand averaged FFRs are shown in Figure 3 for Manipuri group (panel 1), musically trained group (panel 2) and non-tonal group (panel 3) for falling followed by flat and rising tone. As per the Figure 3, the Manipuri group shows clear dark bands of phase-locked activity at f0 and its multiples in response to the Manipuri tonal contours. The non-tonal group shows that the bands are less distinct and more diffuse across all the tones, while musically trained group showed clearer and darker bands than non-tonal group but lesser than the Manipuri group.

Pitch tracking accuracy of Manipuri tones

FFR pitch tracking accuracy was measured by correlating the response contours in comparison to the stimulus contours for all the three groups. The mean pitch tracking accuracy for three Manipuri tones across the groups were plotted in the Figure 4. As per the Figure 4, pitch tracking accuracy was higher for the Manipuri speaking group as compared to the non-tonal language group and musically trained group.

A Mixed ANOVA on pitch tracking accuracy was performed to assess significant effects of the three Manipuri tonal contours across the three groups of subjects. The analysis revealed, significant main effects of tone [$F(2, 42) = 14.7$, $p<0.001$] and group [$F(2, 21) = 12.7$, $p<0.001$]. The group x tone interaction effects were not significant [$F(2, 21) = 2.4$, $p = 0.11$]. A separate one way ANOVA was carried out to see if the mean difference reaches significance for each of the three tones.

Pitch strength of Manipuri tones

FFR pitch strength was measured by the average magnitude of the normalized autocorrelation peak per group for six tonal sections (5-45, 45-85, 85-125, 125-165, 165-205, 205-245 ms) of 40 ms each, for all the three Manipuri tones and were analyzed. The mean pitch strength for all the groups and each of the three tones i.e. falling, rising and flat are depicted in the Figure 5. As per the Figure 5, the overall pitch strength was found to be greater for the Manipuri group as compared to the non-tonal group.

It can be noted from the Figure 5, for the falling tone, the pitch strength was higher across all the sections for the Manipuri group than the other two groups except at section 5 to 45 ms. For the rising tone, the pitch strength of the Manipuri group and musically trained group was greater than the non-tonal group at all sections except 5 to 45 ms, 45 to 85 ms and 85 to 125 ms. At section 205 to 245 ms, the musically trained group showed lower pitch strength than the Manipuri group. For the flat tone, there was minimal difference in the pitch strength for all the groups except at sections 165 to 205 ms and 205 to 245 ms whereas the Manipuri group and musically trained group showed a better pitch strength than the non- tonal group.

A mixed ANOVA was performed to assess the significant effects of the three Manipuri tones across the three groups of subjects. For falling tone, it revealed significant main effects of group [F (2, 27) = 3.33, p <0.01], sections [F (5, 27) = 5.92, p <0.01] but the group x sections interaction [F (10, 27) =0.97, p = 0.47] was not significant. A one way ANOVA was performed to assess whether the mean difference in pitch strength reaches significance across the groups at each section separately.

The analysis revealed a significant difference across groups for all the sections except at sections 5to 45 ms and 85 to 125 ms. Bonferroni's post-hoc analysis revealed a significant mean difference in pitch strength between Manipuri group and non-tonal group at all the sections except at 5to 45 ms and 85 to 125 ms.

For rising tone mixed ANOVA revealed, no significant main effects of group [F (2,27) =1.44, p = 0.25] and sections [F (5,27) = 1.91, p = 0.09], but the interaction group x section was significant [F (10,27) = 2.32, p <0.01]. Tosee the interaction effect, a one way ANOVA was performed, it revealed a significant difference at section 165 to 205 ms between Manipuri group and

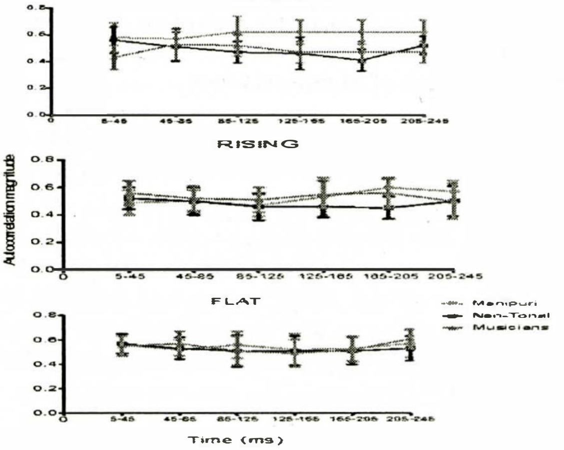

Figure 5: Mean pitch strength of the Manipuri tones within each of their six 40ms sections falling (Panel 1), rising (Panel 2) and flat (Panel 3) as derived from the averaged FFR waveforms of the three groups (Manipuri, Non-tonal language, musically trained).

non-tonal group (p<0.05). For flat tone mixed ANOVA revealed, no significant main effects of group [F (2,27) = 0.55, p = 0.58], sections [F (5,27) = 0.17, p = 0.31] and group x tone interaction [F (10, 27) = 0.34, p = 0.21]. As there was no main effect of group, sections and their interaction, no further analysis was done. Figure 6. provides a better visualization of the one way ANOVA results, showings at which sections there is a significant difference across the groups for all three Manipuri tones.

Figure 6:*The results of one way ANOVA for the three Manipuri tones, falling (column 1), flat (column 2) and rising (column 3) comparing the mean pitch strength between Manipuri group and non-tonal group (row 1) and Manipuri group and musically trained group (row2). The darkened sections show a significant difference amongst the compared groups.*

The results of the present study showed that the pitch tracking accuracy for falling tone was statistically higher in the Manipuri group and musically trained group than the non-tonal group. Of these, Manipuri group showed most accurate pitch tracking compared to other groups. No statistical difference in the pitch tracking accuracy was found across all the groups for the rising and the flat tone. On whole, the pitch strength was more robust for the Manipuri group than the other groups. There was no statistical significant difference across all the subjects for any combination of groups.

For the falling tone, there was significant difference between Manipuri group and non-tonal group at sections 45to 85 ms, 125 to 165 ms, 165 to 205 ms, 205 to 245 ms, but between Manipuri group and musically trained group, only at sections 165 to 205 ms and 205 to 245 ms. Whereas for the rising tone, a significant difference was seen across all the groups only at section 165 to 205 ms. For the flat tone no significant difference was seen in any section across the groups.

Discussion

To accomplish the aim of the present study, brainstem response were obtained using FFRs for three DyIRN, each representing three different tonal contours falling, rising and flat. The results demonstrate that the Manipuri group had greater pitch tracking accuracy and pitch strength than the non-tonal groups. It is further noted that, the falling tone and the rising tone showed more difference for the subcortical pitch representation between tonal group and non-tonal group, whereas for the flat tone, the difference was minimal. Similar to the present study, Krishnan, Gandour, Bidelman, &Swaminathan, (2009) conducted a study comparing the subcortical pitch representation by obtaining FFRs using DyIRN stimuli in tonal (Mandarin) and non-tonal (English) language groups. They also observed similar results as in the present study. However, amount of pitch

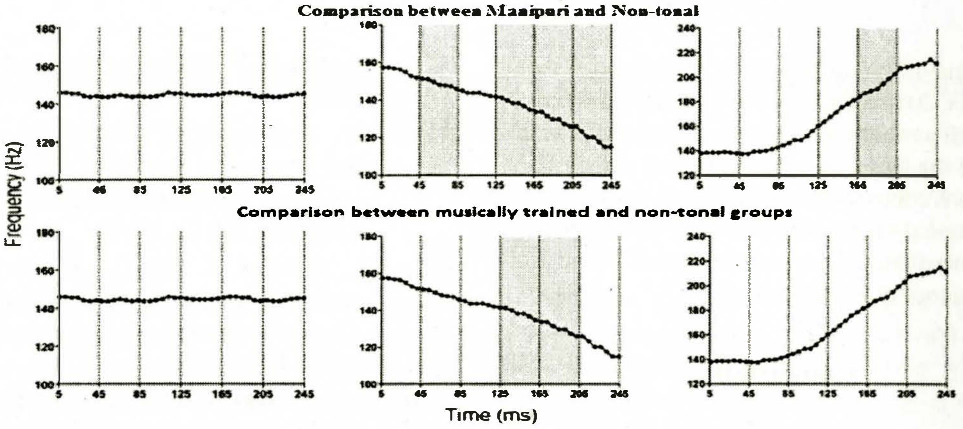

Figure 6: The results of one way ANOVA for the three Manipuri tones, falling (column 1), flat (column 2) and rising (column 3) comparing the mean pitch strength between Manipuri group and non-tonal group (row 1) and Manipuri group and musically trained group (row2). The darkened sections show a significant difference amongst the compared groups.

strength noted in their study was approximately 0.8, whereas in the present study it was 0.6 to 0.7. The precise reasons for reduced pitch strength in the present study are not known. The probable reasons could be the methodological differences, and the instrument used for recording responses. One important difference is the recording system used. The current study used Intelligent Hearing Systems (IHS) Smart EP (Version 4.00) which records at a sampling rate of 16000/sec and averaged waveform has a total of 512 samples. Krishnan et al., (2009) used Tucker-Davis technologies system at a sampling rate of 25000/s and gives 6520 sample for 250 ms.As the number of samples were lesser in the present study, it could have led to a less accurate autocorrelation coefficient value.

In addition, several studies conducted to obtain FFRs using natural or synthetic speech stimuli, to study the subcortical representation of pitch in tonal and non-tonal language speakers, demonstrated similar findings as in the present study (Krishnanet al., 2005; Krishnan et al., 2010).

The results of the present study and previous studies continuously show that linguistically relevant pitch patterns are processed at the brainstem level. The listener's native language experience has changed the way they process linguistically relevant pitch patterns regardless of the stimulus context (non-speech) in which these patterns are embedded. Although the basis for cross-language differences in FFR pitch extraction may emerge from language experience (Manipuri), the effects of such experience are not specific to speech perception (Xu et al., 2006), nor is pitch extraction at the brainstem level necessarily specific to the domain in which pitch patterns occur.

Further it was found that the pitch tracking accuracy in the musically trained group and Manipuri group was not statistically different for the falling tone and the pitch strength for the rising and flat tone. Wong et al., (2007) compared the subcortical pitch representation in terms of pitch tracking accuracy and pitch strength between musicians and non-musicians who were non-native speakers of any tonal language. It was found that the musicians exhibit more faithful pitch tracking and robust encoding of the Mandarin tones than non musicians. This was majorly seen for the falling tone followed by the rising tone. In support to this, the studies conducted to compare the pitch contours with cortical auditory evoked potentials also show similar results (Chandrasekaran, Krishnan, &Gandour, 2009; Fujioka et al., 2004). Further, Song, Skoe, Wong and Kraus (2008) studied the effect of short term linguistic training of pitch tracking accuracy. They noted that there was a decreased number of pitch tracking errors and these decrease was more for the falling contour than the other tonal contours. It can be therefore noted that, along with language experience the long term musical experience also modulates the brainstem processing of musically relevant pitch contours (Munte, Altenmuller, & Jancke, 2002).

The results of the present study and those of previous studies noted that subjects with long term language or musical training have good ability in pitch tracking. Therefore it can be hypothesized that the subcortical pitch processing is shaped by long-term acoustical experience rather than just language experience. That is experience dependent enhancement of pitch representation at the brainstem level is specific to pitch patterns that occur in the listener's acoustical experience. The hypothesis is further supported by studies conducted on comparing two different tonal language groups (Mandarin and Thai) and a non tonal language group (English) using FFR, which reported higher pitch tracking abilities and more robust pitch strength in both of the tonal languages regardless of their language identity, while poorer in non-tonal group (Krishnan, Gandour, &Bidelman, 2010). This implies that the brainstem neurons are differentially sensitive to changes in pitch irrespective of the language identity as long as the two languages have comparable phonological system. Hence the brainstem demonstrates plasticity and the pitch extraction brainstem neurons can be shaped by the acoustical experience of an individual.

Conclusions

From the present study it can be concluded that, FFR can serve as an index of studying subcortical representation of pitch in humans. It can be inferred from the findings that the neural mechanisms underlying the brainstem are adaptive and driven by long-term experience. Over a period of time, they sharpen the pitch extraction neurons for processing pitch contours that are relevant to their language or musical experience. Thus, the brainstem plasticity and pitch representation is the outcome of leaning experience and is not specific to any domain. The study threw light on the nature of subcortical pitch processing in native, non-native and musically trained subjects and the experience dependent plasticity which has implications for future research and understanding the nature of pitch encoding in the human auditory system. . As the results obtained in the present study are in line with most of the research done on subcortical representation of the pitch as reflected by pitch tracking accuracy and pitch strength. It indicates that the both long term and short term learning experience in context of both music and language and is not restricted to either domain.

References

Chandrasekaran, B., Krishnan, A., &Gandour, J. T. (2009b). Sensory Processing of Linguistic Pitch as Reflected by the Mismatch Negativity. *Ear & Hearing, 30* (5), 552-558.

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

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

Griffiths, T. D., Uppenkamp, S., Johnsrude, I., Josephs, O., & Patterson, R. D. (2001). Encoding of the temporal regularity of sound in the human brainstem. *NatureNeuroscience, 4*(6), 633-637.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams :a framework for understanding aspects of the functional anatomy of language. *Cognition, 92*(12), 67-99.

Kraus, N., & Banai, K. (2007). Auditory-processing malleability: Focus on language and music. *Current Direction in Psychological Science, 16*(2), 105-110.

Krishnan, A., Swaminathan, J., & Gandour, J. T. (2009). Experience dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. *Journal of Cognitive Neuroscience, 21*(6), 1092-1105.

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

Krishnan, A., Gandour, J., & Bidelman, G.M. (2010). Effect of tone language on pitch processing in the Brainstem. *Journal of Neurolinguistics, 23,* 81-95.

Munte, T. F.,Altenmuller, E., & Jancke, L.(2002).The musician's brain as a model of neuroplasticity. *Nature Reviews Neuroscience, 3,*473-478.

Song, J.H,.Â Skoe, E., Wong, P.C, & Kraus, N.(2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of cognitive neuroscience. 20,* 1892-1902.

Scott, S. K., & Johnsrude, I. S.,(2003). The neuroanatomical and functional organization of speech perception. *Trends in neurosciences, 26*(2), 100-107.

Swaminathan, J., Krishnan, A., &Gandour, J. T. (2008). Pitch encoding in speech and non-speech contexts in the human auditory brainstem. *Neuroreport, 19*(11), 1163-1167.

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

Xu, Y., Gandour, J. T., & Francis, A. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *Journal of the Acoustical Society of America, 120*(2), 1063-1074.

Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science, 256,* 846-849.