Brainstem Correlates of Speech Perception in Noise: Carnatic Musicians Vs. Non-Musicians

[1]Zubin Vinod & [2] Rajalakshmi K.

Abstract

Many studies have indicated the presence of superior auditory capabilities as a result of long-term musical experience, including better perception of speech in a background of noise. Musicians have life long experience parsing melodies from background harmonies, which can be considered a process analogous to speech perception in noise. To investigate the effect of musical experience on the neural representation of speech-in-noise, the subcortical neurophysiological responses to speech in quiet and noise in a group of highly trained musicians and nonmusician controls were compared. Musicians were found to have a more robust subcortical representation of the acoustic stimulus in the presence of noise. Specifically, musicians demonstrated earlier latencies and higher amplitudes of onset and transition peaks, higher amplitudes of encoded formants and less degraded response morphology in noise. Neural measures were associated with better behavioral performance on the test of Speech Perception in Noise (SPIN) for which musicians outperformed the nonmusician controls. These findings suggest that musical experience limits the negative effects of competing background noise, thereby providing the first biological evidence for musicians' perceptual advantage for speech-in-noise.

Keywords: *Music, Carnatic, Speech ABR, FFR, SPIN*

Introduction

The domains of music and language share many features, the most direct being that both exploit changes in pitch patterns to convey information. Music uses pitch contours and intervals to communicate melodies and tone centers. Pitch patterns in speech convey prosodic information; listeners use prosodic cues to identify indexical information, i.e., information about the speaker's intention as well as emotion and other social factors.

Through years of sensory-motor training, often beginning in early childhood, musicians develop an expertise in their instrument of specialization or mastery over their voice. In the course of training, musicians increasingly learn to attend to the fine-grained acoustics of musical sounds. Attention to pitch, timing and timbre is emphasized during music training. A variety of studies have found that musical training improves auditory-perceptual skills resulting in enhanced behavioural (Jeon & Fricke, 1997; Koelsch, Schroger & Tervaniemi, 1999; Micheyl Delhommeau, Perrot & Oxenham, 2006; Rammsayer & Altenmuller, 2006; Tervaniemi, et al., 2009) and neurophysiological (Brattico, Naatanen & Tervaniemi, 2001; Pantev et al., 2001; Schneider, et al., 2002; Shahin , Bosnyak , Trainor, Roberts & Larrey, 2003; Tervaniemi, et al., 2005; Kuriki, Kanda, & Hirata, 2006; Kraus, Skoe, Parbery-Clark & Ashley, 2009) responses.

It is only reasonable to assume that the benefits that musicians have in processing music would also extend to speech stimuli. A number of research studies have shown that music training benefits auditory processing not only in the musical domain, but also in the processing of speech stimuli (Musacchia et al., 2007; Schon, Magne & Besson, 2004; Wong, Skoe, Russo, Dees & Kraus, 2007). Other verbal and non-verbal skills such as auditory attention (Strait, Kraus, Parbery-Clark, & Ashley, 2010), auditory stream segregation (Beauvois & Meddis, 1997), processing emotion in speech (Strait, Kraus, Skoe & Ashley, 2009), working memory (Chan, Ho & Cheung, 1998; Forgeard, Winner, Norton & Schlaug, 2008) and processing of prosody and linguistic features in speech (Chandrasekaran, Krishnan & Gandour, 2009; Wong, Skoe, Russo, Dees & Kraus, 2007).

Of special note is the enhanced ability of musicians to extract relevant signals from a complex soundscape (e.g., the sound of their own instrument in an orchestra). Speech perception in noise is a complex task that requires the segregation of target signals from a competing background noise. To complicate matters, the noise also degrades the signal particularly by disrupting the perception of rapid spectro-temporal changes (Brandt & Rosen, 1980). Poor performance in the task of speech perception in noise is seen in individuals with hearing impairment (Gordon- Salant & Fitzgibbons, 2004) and language-based learning disabilities (Bradlow, Kraus & Hayes, 2003; Ziegler, Pech-Georgel, George & Lorenzi, 2005) whereas musicians demonstrate better performance than non-musicians (Parbery-Clark, Skoe & Kraus, 2009). It was hypothesized that a musician's long-term experience with musical stream segregation would transfer to the homologous task of speech perception in noise. Parbery-Clark et al. (2009) found a distinct speech in noise advantage for musicians, as measured by two standardized

[1]Email: zforce432@gmail.com,
[2]professor of Audiology, Email: veenasrijaya@gmail.com

tests of hearing in noise (HINT, Hearing in-noise test; QuickSIN). Musicians showed superior working memory and performed better on a frequency discrimination task. Across all participants, the number of years of consistent practice with a musical instrument correlated strongly with performance on QuickSIN, auditory working memory and frequency discrimination. These correlations strongly suggest that practice fine tunes cognitive and sensory ability, leading to an overall advantage in speech perception in noise in musicians.

All these enhanced abilities in musicians may be related to structural and functional enhancements seen at different levels of their nervous system. For instance, musicians have more neural cell bodies (grey matter volume) in the auditory, motor and visuo-spatial areas of the brain (Gaser & Schlaug, 2003) and also have more axonal projections that connect the right and left hemispheres (Schlaug, Jancke, Huang, Staiger & Steinmetz, 1995). All these anatomical enhancements are seen to translate into improved auditory and cognitive skills as is evidenced by various studies. The intensive practice over the years has been attributed to bring about neuroplastic changes in the practitioner as is evidenced in many research studies (Pantev et al., 1998; Koelsch et al., 1999; Pantev, Roberts, Schulz, Engelien & Ross, 2001; Tervaniemi, Rytkonen, Schroger, Ilmoniemi & Naatanen, 2001; Fujioka, Trainor, Ross, Kakigi & Pantev, 2005; Musacchia, Sams, Skoe & Kraus, 2007;). One of the mechanisms used to explain the findings of music-induced experience dependent plasticity at the level of the brainstem is increased efficiency of top-down predictive coding (Strait et al., 2010). Recent studies have suggested an important role for the feedback (top-down) pathways in fine-tuning the auditory signal at early stages of auditory processing (Luo, Wang, Kashani & Yan, 2008). Such top-down influences back-project all the way to the cochlea through the medial olivocochlear bundle (MOCB). These authors have said that feedback initiated by the higher (cortical) structures is transferred to the lower (brainstem) structures via the efferent auditory system. This results in an enhanced selectivity of sound features at the lowest levels of the auditory system which is important for higher-level structures to distinguish relevant information in the signal from irrelevant details. The human auditory brainstem response (ABR) has been used as an index of brainstem encoding of speech stimuli (Chandrasekaran & Kraus, 2010; Skoe & Kraus, 2010).

Since the FFR preserves spectral information up to about ~2000 Hz and reflects neural timing in the order of milliseconds, it can therefore be used to examine the fidelity of the brainstem representation of spectral and timing information. It has been found that the addition of background noise delays the timing of brainstem responses (Cunningham, Nicol, Zecker & Kraus, 2000; Cunningham, Nicol, Zecker, Bradlow & Kraus, 2001) and reduces spectral magnitude. There is evidence from studies using speech-evoked ABR that music training modulates the effect of background noise on subcortical auditory representation (Parbery- Clark, et al., 2009). Musicians show less degraded brainstem representation of speech relative to non-musicians, as evidenced by faster neural timing, enhanced spectral representation, and better stimulus-to-response correlations. Though the differences between musicians and non-musicians are present even in quiet backgrounds (Musacchia et al., 2007), it is in the presence of background noise that the differences in spectral representation between musicians and non-musicians are large, suggesting that musical experience protects against the debilitating effects of background noise (Parbery-Clark, et al., 2009). Thus timing and spectral features are preserved at the level of the brainstem to a greater extent due to musical experience and these enhancements translate into a better performance on the task of speech perception in noise. The speech-evoked ABR is hence considered to be a reliable indicator of the biological basis of speech perception in noise.

Despite the considerable amount of literature dealing with the enhanced subcortical encoding of speech in the presence of noise in Western musicians, there is a dearth of similar studies in Carnatic musicians. Thus the following study was carried out to verify whether trained Carnatic musicians show better perception of speech in the presence of background noise as compared to non-musicians and if so, whether they had enhanced subcortical encoding of speech stimuli as measured via speech evoked ABR as compared to non-musicans.

Method

Subjects

Fifteen musicians and fifteen nonmusicians participated in this study. Participants' age ranged from 18 to 30 years. Participants categorized as musicians started training in Carnatic music before the age of 8 and practiced consistently for at least10 years before enrolling in the study. Nonmusicians were required to have had no musical training. All participants had normal hearing thresholds from 125 to 8000 Hz. No participant reported any cognitive or neurological deficits.

Stimuli

The /da/ stimulus is a 40 ms synthesized speech syllable produced using KLATT synthesizer (Klatt, 1980). This stimulus simultaneously contains the broad spectral and fast temporal information characteristic of stop consonants, and spectrally rich formant transitions between the consonant and the steady-state vowel. Although the steady-state portion is not present, the stimulus is still perceived as being a consonant-vowel syllable. The fundamental frequency (F0) linearly rises from

103 to 125 Hz with voicing beginning at 5 ms and an onset noise burst during the first 10 ms. 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. The phonemically balanced wordlist in Kannada (Yathiraj & Vijayalakshmi, 2005) was presented in the presence of ipsilateral speech noise to assess the patient's perception of speech in noise.

Procedure

The speech syllable /da/ was presented in condensation and rarefaction polarities at 80 dB sound pressure level (SPL) through insert ear phones (ER-3; Etymotic Research). In the noise condition, both the /da/ and white noise were presented simultaneously to the test ear. The /da/ was presented at a 0 dB signal-to-noise ratio over the background noise.

The responses to two background conditions, quiet and noise, were collected using Bio-Logic Navigator Pro EP with 3 gold disc electrodes which were fastened to the scalp. Responses were differentially recorded with a vertical montage (Cz active, forehead ground, and earlobe references), an optimal montage for recording brainstem activity (Galbraith et al., 1995; Chandrasekaran & Kraus, 2009). Contact impedance was 2 kΩ or less between electrodes. Three thousand artifact-free sweeps were recorded for each condition for both polarities. Participants were asked to sleep for the recording session. To limit the inclusion of low-frequency cortical activity, brainstem responses were off-line bandpass filtered from 70 to 2000 Hz (12 dB/octave, zero phase-shift) using Bio-logic Navigator Pro EP. The filtered recordings were epoched using a time window of 64 ms which included a prestimulus time of 10 ms (default setting in Biologic system) with the stimulus onset occurring at 0 ms. Any sweep with activity greater than 35μV was considered artifact and rejected. The responses to the two polarities were added together to minimize the presence of the cochlear microphonic and stimulus artifact on the neural response (Gorga et al., 1985; Aiken & Picton, 2008). Last, responses were amplitude-baselined to the prestimulus period.

Analysis

The latency and amplitude of onset peak V and transition peaks D, E and F were measured. The waveforms obtained in both conditions were also Fast Fourier Transform (FFT) to obtain information regarding the spectral characteristics of the FFR (frequency and amplitude of spectral peaks). The average spectral amplitude was calculated for a frequency range from 103 to 120 Hz which encompasses the fundamental frequency (F0). FFT was performed on all speech evoked potentials using a custom made program run in MATLAB. The peak amplitude corresponding to F0 was also calculated using a custom made program file in the MATLAB platform. The frequency analysis was done from 11.4 to 40.6 ms. The sustained portion of the response (FFR) was passed through 103 to 120Hz band pass fourth order Butterworth filters in order to obtain the energy at F0. The Fourier analysis was then performed on the filtered signal. A subject's responses were required to be above the noise floor in order to be included in the analysis. This was performed by comparing the spectral magnitude of pre stimulus period to that of the response. If the quotient of the magnitude of F0 frequency component of FFR divided by the pre stimulus period was >1, the response was deemed to be above the noise floor. Statistical analyses were done using SPSS 20.

Results

In quiet, the onset peaks V and the transition peaks D, E, and F were clearly visible in the speech evoked ABR of the non-musicians. The morphology of the waves was noticeably poorer in noise, with peaks having reduced amplitude and delayed latencies. As in Figure 1, the V-A complex is almost eliminated in noise, though the transition waves are less affected. In quiet, the morphology of the Speech-Evoked ABR of musicians did not vary much from that seen in non-musicians.

Though the waveform morphology was poorer in noise than in quiet, the waves were by and large better defined than in the corresponding waveforms of non-musicians. The V-A complex in particular is more clearly seen (Figure 2).

Comparison of Peak Latencies

The comparison of latency measures obtained in different conditions across the groups using mixed ANOVA reveals the presence of main effects of conditions [$F_{(1,28)} = 115.146$, $p<0.001$] and groups[$F_{(1,28)} = 27.664$, $p<0.001$] as well as interaction effects between conditions and groups[$F_{(1,28)} = 10.745$, $p=0.003$], latency and groups[$F_{(3,84)} = 20019.337$, $p<0.001$], conditions and latency [$F_{(3,84)} = 9.087$, $p= 0.022$] and conditions,

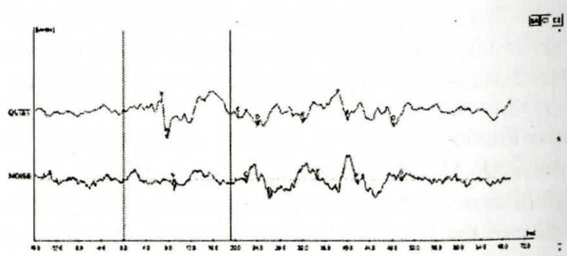

Figure 1: Speech-Evoked ABR in response to 40 ms /da/ acquired in a Non-Musician in quiet and in noise (0dB SNR).

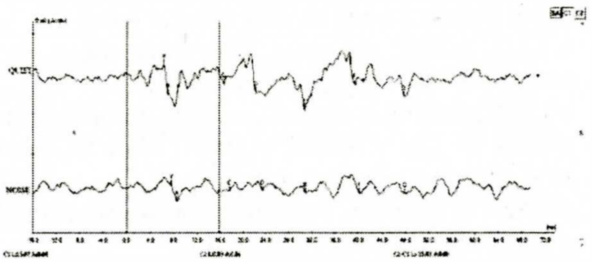

Figure 2: Speech-Evoked ABR in response to 40 m /da/ acquired in a Musician in quiet and in noise (0dB SNR).

latencies and groups [$F_{(3,84)}$ = 3.389, p=0.005].

Descriptive statistics were also done to find out the mean and standard deviation of the latencies for musicians and non-musicians in quiet and in noise. Paired t-test was also carried out to check for the presence of significant differences in the latencies of the waves acquired in quiet and noise in each group.

It is evident from the results that musicians showed earlier mean latencies of all the waves than non-musicians in quiet (Table 1) and in noise (Table 2).

There are also significant differences in the latencies of wave V [$t_{(14)}$= -9.909, p<0.001], wave D [$t_{(14)}$= -6.633, p<0.001], wave E [$t_{(14)}$= -6.859, p<0.001] and Wave F [$t_{(14)}$= -5.135, p<0.001]in quiet and in noise in non-musicians. The latencies of the waves in quiet and in noise were also found to be significantly different for musicians for wave V [$t_{(14)}$= -8.006, p<0.001], wave D [$t_{(14)}$= -3.938, p<0.001], wave E [$t_{(14)}$= -5.121, p<0.001] and wave F [$t_{(14)}$= -7.371, p<0.001].

Table 1: Mean and Standard Deviation for Peak Latencies (in milliseconds) of Non-Musicians and Musicians in Quiet

Measure	Group	**Mean**	Standard Deviation
Wave V Latency	Nonmusicians	6.23	0.32
	Musicians	6.02	0.18
Wave D Latency	Nonmusicians	22.88	0.72
	Musicians	22.01	0.34
Wave E Latency	Nonmusicians	31.04	0.61
	Musicians	30.42	0.32
Wave F Latency	Nonmusicians	39.54	0.66
	Musicians	38.97	0.23

Table 2: Mean and Standard Deviation for Peak Latencies (in milliseconds) of Non-Musicians and Musicians in Noise

Measure	Group	Mean	Standard Deviation
Wave V Latency	Nonmusicians	7.68	0.71
	Musicians	7.43	0.74
Wave D Latency	Nonmusicians	25.35	1.50
	Musicians	22.96	0.91
Wave E Latency	Nonmusicians	33.80	1.65
	Musicians	31.97	1.08
Wave F Latency	Nonmusicians	41.96	2.05
	Musicians	39.89	0.63

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

The results of the one way MANOVA show that in quiet, the latencies of waveV [$F_{(1,28)}$ = 4.725, p=0.038], wave D [$F_{(1,28)}$ = 17.535, p<0.001], Wave E[$F_{(1,28)}$ = 12.165, p=0.002] and Wave F[$F_{(1,28)}$ = 9.684, p=0.004] were found to be significantly different across the groups. In the presence of noise, the latency of wave V [$F_{(1,28)}$ = 0.890, p=0.353] did not differ significantly across the groups but the latencies of the D [$F_{(1,28)}$ = 27.614, p<0.001], E [$F_{(1,28)}$ = 12.774, p=0.001] and F [$F_{(1,28)}$ = 13.979, p=0.001] did vary.

Comparison of Amplitude Measures

Mixed ANOVA was carried out to compare between the groups for amplitude measures obtained the conditions of quiet and in noise (Table 8). The results of the test reveals the presence of main effects of conditions [$F_{(1,28)}$ = 576.733, p<0.001] and groups[$F_{(1,28)}$ = 19.332, p<0.001] as well as interaction effects between conditions and groups[$F_{(1,28)}$ = 14.248, p=0.001] , amplitude and groups[$F_{(3,84)}$ = 24.940, p<0.001], conditions and latency [$F_{(3,84)}$ = 9.969, p<0.001]and conditions, latencies and groups [$F_{(3,84)}$ = 53.356, p<0.001].

Descriptive statistics were done to find out the mean and standard deviation of the latency measures for musicians and non-musicians in quiet and in noise. The examination of the mean amplitudes of the waves V, D, E and F reveals that the musicians had higher mean amplitudes than non- musicians for all the waves in quiet (Table 3). The mean amplitudes of all the waves acquired in noise were also greater for musicians than for non-musicians (Table 4).

Paired t-test was also carried out to compare the amplitudes of the waves acquired in quiet and noise in each group. In non-musicians, the amplitudes of wave V [$t_{(14)}$= 10.505, p<0.001], wave D [$t_{(14)}$= 5.922, p<0.001], wave E [$t_{(14)}$= 7.388, p<0.001] and wave F [$t_{(14)}$= 7.542, p<0.001] were found to be significantly greater in quiet than in noise. The amplitudes of wave V [$t_{(14)}$= 14.086, p<0.001], wave D [$t_{(14)}$= 15.282, p<0.001], wave E [$t_{(14)}$= 14.826, p<0.001] and wave F [$t_{(14)}$= 3.967, p<0.001] were also found to be significantly greater in quiet than in noise in musicians.

To compare the amplitude measures of the different waves across the 2 groups in quiet and in noise, two

Table 3: Mean and Standard Deviation for Peak Amplitudes (in micro Volts) of Non-Musicians and Musicians in Quiet

Parameter	Group	Mean (μV)	Standard Deviation
Amplitude of peak V	Nonmusicians	0.22	0.049
	Musicians	0.27	0.050
Amplitude of peak D	Nonmusicians	0.15	0.073
	Musicians	0.18	0.034
Amplitude of peak E	Nonmusicians	0.22	0.091
	Musicians	0.41	0.097
Amplitude of peak F	Nonmusicians	0.16	0.049
	Musicians	0.17	0.072

Table 4: Mean and Standard Deviation for Peak Amplitudes (in micro volts) of Non-Musicians and Musicians in Noise

Parameter	Group	Mean (μV)	Standard Deviation
Wave V Amplitude	Nonmusicians	0.043	0.031
	Musicians	0.078	0.023
Wave D Amplitude	Nonmusicians	0.050	0.025
	Musicians	0.067	0.015
Wave E Amplitude	Nonmusicians	0.044	0.027
	Musicians	0.056	0.014
Wave F Amplitude	Nonmusicians	0.087	0.053
	Musicians	0.105	0.035

measures of one-way MANOVA were carried out. The results show that in quiet, the amplitudes of the waves V [$F_{(1,28)} = 8.196$, $p=0.008$]and E [$F_{(1,28)} = 29.932$, $p=<0.001$] are significantly greater in musicians than in non-musicians but the amplitudes of wave D [$F_{(1,28)} = 2.130$, $p=0.156$] and wave F [$F_{(1,28)} = 0.330$, $p=0.570$] were not. In noise, the amplitudes of waves V [$F_{(1,28)} = 12.078$, $p=0.002$] and D [$F_{(1,28)} = 4.709$, $p=0.039$] were found to be significantly greater in musicians than in non-musicians. However the differences n amplitudes was not the significant for wave E [$F_{(1,28)} = 2.461$, $p=0.128$] and wave F [$F_{(1,28)} = 1.172$, $p=0.288$].

Comparison of Formant Amplitude Measures

Mixed ANOVA was carried out to compare between the groups for formant amplitude measures obtained the conditions of quiet and in noise. The results indicated as to the presence of a main effect of condition [$F_{(1,28)} = 0.330$, $p<0.001$] while no main effect of group [$F_{(1,28)} = 3.328$, $p=.079$] was noted. Interaction effects were present between formant amplitudes & groups [$F_{(2,56)} = 3.723$, $p=0.030$] as well as conditions & formant amplitudes [$F_{(2,56)} = 64.899$, $p<0.001$] while no significant effects of condition and group [$F_{(1,28)} = 0.017$, $p=0.898$] & condition, formant amplitude and group [$F_{(2,56)} = 0.061$, $p=0.941$] were noticed.

Descriptive statistics were carried out to find out the mean and standard deviation of the formant amplitude measures for musicians and non-musicians in quiet and in noise. The mean values of the amplitudes of formants F0, F1 and F2 were found to be greater for musicians than non-musicians in quiet (Table 5).The mean formant amplitudes were higher for musicians than non-musicians in waveforms acquired in noise (Table 6).

Paired t-test was also carried out to compare the amplitudes of the formants in quiet and noise in each group. In non-musicians, the amplitudes of F0 [$t_{(14)}= 6.943$, $p<0.001$], F1 [$t_{(14)}= 6.872$, $p<0.001$] and F2 [$t_{(14)}= 0.684$, $p<0.001$] were seen to be significantly greater in quiet than in noise. In musicians, the amplitudes of F0 [$t_{(14)}= 5.386$, $p<0.001$], F1 [$t_{(14)}= 5.111$, $p<0.001$] and F2 [$t_{(14)}= 1.128$, $p<0.001$] were seen to be significantly greater in quiet than in noise. The mean values of the formant amplitudes were seen to be lesser in noise than in quiet.

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

The results of the One Way MANOVA tests show that the formant amplitudes of F0 [$F_{(1,28)} = 2.401$, $p=0.132$], Fl [$F_{(1,28)} = 0.141$, $p=0.710$] and F2 [$F_{(1,28)} = 0.104$, $p= 0.750$] were not found to be significantly greater for musician as compared to non musicians and in quiet. In noise, the same trend was observed across both the groups for amplitudes of F0 [$F_{(1,28)} = 2.636$, $p=0.116$], Fl [$F_{(1,28)} = 1.477$, $p=0.234$] and F2 [$F_{(1,28)} = 0.013$, $p=0.908$].

Table 5: Mean and Standard Deviation for Formant Amplitudes (in dB) of Non-Musicians and Musicians in Quiet

Parameter	Group	Mean(dB)	Standard Deviation
Amplitude of F0	Nonmusicians	5.10	1.64
	Musicians	6.16	2.05
Amplitude of F1	Nonmusicians	0.62	0.20
	Musicians	0.65	0.24
Amplitude of F2	Nonmusicians	0.19	0.049
	Musicians	0.19	0.063

Table 6: Mean and Standard Deviation for Formant Amplitudes (in dB) of Non-Musicians and Musicians in Noise

Parameter	Group	Mean (dB)	Standard Deviation
Amplitude of F0	Nonmusicians	2.23	1.31
	Musicians	3.14	1.73
Amplitude of F1	Nonmusicians	0.25	0.11
	Musicians	0.33	0.23
Amplitude of F2	Nonmusicians	0.17	0.08
	Musicians	0.17	0.05

Table 7: Results of Independent Sample t-Test for comparing SPIN Scores across the 2 groups

Parameter	| t |	Degrees of Freedom	p(2-tailed)
SPIN scores	-3.500	28	0.002

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

Table 8: Mean and Standard Deviation of SPIN Score in Musicians and Non-Musicians

Group	Mean	Standard Deviation
Non- Musicians	78.667	3.266
Musicians	82.400	2.529

Comparison of SPIN Scores

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

Table 9: Results of Pearson's Correlation: Correlation of SPIN Scores with Latency, Amplitude and Formant Amplitude in Quiet for Non-Musicians

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

**The mean difference is significant at the 0.05 level.*

Table 10: Results of Pearson's Correlation: Correlation of SPIN Scores with Latency, Amplitude and Formant Amplitude in Quiet for Musicians

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

**The mean difference is significant at the 0.05 level.*

An independent sample t-test was carried out to compare SPIN scores across the 2 groups. The results indicate that the scores differ significantly across the 2 groups (Table 7). It may be seen from Table 8 that the musicians had a higher mean score on the SPIN test than the non-musicians.

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

In non-musicians, the SPIN scores did not correlate with any of the measures obtained in quiet (Table 9). In musicians, the SPIN scores were found to correlate negatively with the latencies of waves D and F obtained in quiet (Table 10).

In noise, it was seen that for non-musicians, the SPIN scores negatively correlated with the latency of wave V and positively correlated with the amplitudes of wave V and D (Table 11).

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

In noise, the SPIN scores of musicians correlated neg-

Table 11: Results of Pearson's Correlation: Correlation of SPIN Scores with Latency, Amplitude and Formant Amplitude in Noise for Non-Musicians

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

**The mean difference is significant at the 0.05 level.*

Table 12: Results of Pearson's Correlation: Correlation of SPIN Scores with Latency, Amplitude and Formant Amplitude in Noise for Musicians

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

**The mean difference is significant at the 0.05 level.*

atively with the latencies of wave V and F (Table 12), indicating that subjects with earlier wave V and F latencies showed better performance on the task of speech perception in noise. Positive correlation was seen with the formant amplitude of F0 obtained in noise, indicating that the superior encoding of F0 in musicians had resulted in enhanced SPIN scores.

Discussion

Wave Morphology

It was seen that the addition of noise to the speech stimulus caused the morphology to deteriorate significantly from the quiet condition in both musicians and non-musicians. Similar findings were reported in studies by Russo, Nicol, Zecker, Hayes and Kraus (2004) and Russo, Nicol, Musacchia and Kraus (2004). However, musicians were seen to have a comparatively better morphology of the waveform in the presence of noise than non-musicians. This is in line with the findings of Parbery-Clark, Skoe and Kraus (2009).

Latencies of Onset and Transition Peaks

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

In the present study, in both non-musicians and musicians, the latencies of all the waves were seen to be significantly different in quiet and noise, with delay in latencies of the waves acquired in noise.The above findings are in agreement with Russo, Nicol, Musacchia and Kraus (2004) who documented the detrimental effects of noise on the subcortical representation of speech signals. The same findings were also reported by Parbery-Clark, et al., (2009).

Musicians showed significantly earlier mean latencies of all the waves than non-musicians in quiet. This is in agreement with the findings of Musacchia, et al., (2007) who found that musicians had earlier wave latencies than non-musicians in quiet. Musacchia, Stait and Kraus (2008) also documented the onset timing of musicians in quiet to be earlier than that of non-musicians. However, in contradiction Parbery-Clark, et al., (2009) found that the latencies of the waves were not significantly different in musicians and non-musicians in quiet.

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

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

Amplitudes of Onset and Transition Peaks

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

In the present study, both non-musicians and musicians, the amplitudes of all the waves were seen to be significantly greater in quiet than in noise. This indicates that the noise has a detrimental effect on the subcortical representation of the signal (Don & Eggermont, 1978; Cunningham et al., 2001; Russo et al., 2004). Russo, et al., (2004) and Parbery-Clark, et al., (2009) have also documented reduced amplitude of the onset and transition waves in the presence of background noise.

Musicians had higher mean amplitudes than non- musicians for all the waves in quiet, though only the amplitudes of the waves V and E are significantly different across the groups. This finding is in agreement with those of Musacchia, et al., (2007) and Parbery-Clark, et al., (2009). Parbery-Clark, et al., (2009) had documented that there were no significant differences in the amplitudes of the onset and transition waves in quiet across musicians and non-musicians, though the mean amplitudes were found to be greater for musicians.

The mean amplitudes of all the waves acquired in noise were also greater for musicians than for non-musicians, with significant differences seen in the amplitudes of waves V and D. Parbery-Clark, et al., (2009) also documented the reduction in amplitude of the onset and transition peaks in the presence of background noise to be similar in musicians and non-musicians. Though the mean amplitude of the transition wave was found to be greater in musicians, the same had not been observed with the onset wave. However, it may be pointed out that the amplitudes of onset responses are highly variable (Starr & Don, 1988; Hood, 1998) and this fact may have contributed to the differences present between the two studies.

From the above results, it is seen that the musicians have overall higher mean amplitudes of different waves in both quiet and in noise when compared to non-musicians. This is due to the disruption of the neural representation of stimulus features by noise (Russo et al., 2004). However, due to the training musicians undergo which involves the selective attention to a specific element from a complex soundscape, there is an enhanced encoding which improves the subcortical signal quality, resulting in a more robust representation of the target acoustic signal in noise. This once again points to the fact that musical training helps strengthen the sub-cortical representation of the stimulus features via top-down processes.

Formant Amplitudes

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

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

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

From the above findings, it was seen that both groups also showed higher mean formant amplitudes in quiet than in the presence of noise, evidence to the degradation of the neural representation of the speech signal in the presence of noise. The musicians also showed higher mean formant amplitudes than the non-musicians in both quiet and in noise, though the differences were not statistically significant. The enhanced encoding of the formants of the speech stimulus in musicians has been documented by many authors (Musacchia et al., 2007; Wong, et al., 2007). The higher mean formant amplitudes of musicians in noise as compared to non-musicians indicates a more robust sub-cortical representation of the speech signal, possibly brought about by years of continuous musical training. One possible explanation for this finding is the based on the Hebbian principle, which posits that the associations between neurons that are simultaneously active are strengthened and those that are not are subsequently weakened (Hebb, 1949). Given the present results, we can speculate that extensive musical training may lead to greater neural coherence, especially pertaining to relevant features crucial to the identification of the stimulus. This strengthening of the underlying neural circuitry would lead to a better bottom-up, feed-forward representation of the signal. We can also interpret these data within the framework of corticofugal modulation in which cortical processes shape the afferent auditory encoding via top-down processes as mentioned earlier in the discussion. Though we cannot separate the contributions of top-down and bottom-up processing, they are not mutually exclusive explanations. In all likelihood, top-down and bottom-up processes are reciprocally interactive with both contributing to the subcortical changes observed with musical training.

Comparison of Wave Latency, Wave Amplitude and Formant Amplitude Measures with Speech Perception Scores in Noise (SPIN Scores)

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

The SPIN scores differed significantly across the 2 groups. Musicians had a higher mean score on the SPIN test than the non-musicians. Parbery-Clark, Skoe, Lam and Kraus (2009) and Parbery-Clark, et al., (2009) also report of a distinct advantage in musicians on the task of perception of speech in noise. This advantage was reported to correlate well with the number of years of training the musician had undergone, which strongly suggested that such intensive training helps to fine tune sensory and cognitive processes that contributed to the task of speech perception in noise.

Upon investigation as to whether the superior performance of musicians over non-musicians on the task of speech perception in noise was related to the differences in the subcortical encoding of speech stimuli across the two groups, it was found that in non-musicians, the SPIN scores did not correlate with any of the measures obtained in quiet. In musicians, the SPIN scores were found to correlate negatively with the latencies of waves D and F obtained in quiet. However, Parbery-Clark, et al., (2009) found that in quiet, there was no significant correlation between latency, amplitude or formant amplitude of brainstem responses of a subject and the corresponding scores on the task of speech perception in noise.

For waveforms acquired in noise, it was seen that for non-musicians, the SPIN scores negatively correlated with the latency of wave V and positively correlated with the amplitudes of wave V and D. Thus, poorer performance on the SPIN test was found to be related to the prolongation of onset latency and the reduction of amplitudes of the onset wave V and transition wave D in non-musicians, indicating that addition of noise had resulted in disruption of brainstem timing and a reduction in the amplitude of the responses encoding stimulus features (onset and transition), which had resulted in reduced SPIN scores. The SPIN scores of musicians correlated negatively with the latencies of wave V and F, indicating that subjects with earlier wave V and F latencies showed better performance on the task of speech perception in noise. Positive correlation was seen with the formant amplitude of F0 obtained in noise, indicating that the superior encoding of F0 in musicians had resulted in enhanced SPIN scores.

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

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

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

Conclusions

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

References

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

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

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

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

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

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

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

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

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

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

Don, M., & Eggermont, J.J. (1978). Analysis of the click-evoked brainstem potentials in man using high-pass noise masking. *Journal of Acoustic Society of America, 63*, 1084 -1092.

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

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

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.

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

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

Gordon-Salant, S., & Fitzgibbons, P. J., (2004). Effects of stimulus and noise rate variability on speech perception by younger and older adults. *Journal of Acoustic Society of America, 115*, 1808-17.

Gorga, M., Abbas, P., & Worthington, D. (1985). Stimulus calibration in ABR measurements. In: The auditory brainstem response (Jacobsen J, ed), pp 49 -62. San Diego: College-Hill.

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

Hood, L. (1998). Clinical applications of the auditory brainstem response. San Diego: Singular.

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

Klatt, D., (1980) Software for a Cascade/Parallel Formant Synthesizer. *Journal of Acoustic Society of America, 67*, 13-33.

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

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

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

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

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

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

Musacchia, G., Strait, D., Kraus, N. (2008). Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians. *Hearing Research, 24*, 34 -42.

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

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

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

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

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

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

Schneider, P., Scherg, M., Dosch, H. G., Specht, H.J., Gutschalk, A.,Â &Â Rupp,A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Natural Neuroscience, 5*, 688-694.

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

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

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

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

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

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

Starr, A., & Don, M. (1988). Brain potentials evoked by acoustic stimuli. In, human event-related potentials. Handbook of electroencephalography and clinical neurophysiology, Vol 3 (Picton T, ed), pp 97-158. New York: Elsevier.

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

Tervaniemi, M., RytkoÂ¨nen, M., SchroÂ¨ger, E., Ilmoniemi, R. J, & NaÂ¨aÂ¨taÂ¨nen, R. (2001). Superior formation of cortical memory traces for melodic patterns in musicians. *Learning and Memory, 8*, 295-300

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

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

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

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