**The Effect of music TRAINING AND family background of music on temporal resolution abilities and Speech perception in Noise**

Dr. Rajalakshmi K.

Professor of Audiology

Dept. of Audiology

All India Institute of Speech & Hearing,

Mysore-06

Ph no: 9342187872

e-mail: [veenasrijaya@gmail.com](mailto:veenasrijaya@gmail.com)

Bharath Bhushan K.R.

Audiologist

Cadential Technologies Pvt. Ltd

Bangalore-34

Ph no: 9844632911

e-mail: [bharathaudiology@gmail.com](mailto:bharathaudiology@gmail.com)

**The Effect of music TRAINING AND family background of music on temporal resolution abilities and Speech perception in Noise**

***Abstract***

***Objective:*** *The objective of the present study was to study the relationship between temporal resolution and speech perception in noise among children with and without family background of music.*

***Method:*** *Totally 40 subjects aged between 12-16 years participated, 20 subjects formed group ‘A’ who received formal musical training, had a family with background of music and the rest 20 subjects formed group ‘B’ who did not receive any formal musical training and did not have any family with background of music. Both the groups were subjected to Gap Detection Test, Temporal Modulation Transfer Function and Quick Speech in Noise test in Kannada.*

***Results:*** *The results from the present study showed that the children who had formal musical training, who also had a family with musical background exhibited better temporal resolution abilities and ability to perceive speech in the presence of noise when compared to those children without formal musical training and without a musical family background. Musical training as a factor has contributed to better performance whether or otherwise of the family background. However, for those children who also had family background of music along with the formal musical training, the auditory processing skills were shaped to the finest level.*

***Conclusion:*** *Music training leads to better auditory processing abilities. In the context of familial inheritance, the auditory processing abilities get more finely tuned and provides for better performance on the tasks measured in this study.*

***Implication:*** *Music training can be advocated for all children irrespective of whether they have musical background or otherwise to facilitate speech perception in adverse listening conditions and probably to provide the advantage of enhanced auditory working memory.*

***Key words****: Plasticity, gap detection, temporal modulation transfer function,*

**INTRODUCTION**

Music has its presence in every culture all around the world. Earlier researchers have focused on how formal music training impacts various aspects of cognitive development such as auditory perception, memory, and language skills. The present study was aimed at finding the relationship between temporal resolution abilities and speech perception in noise among individuals who have undergone formal musical training and individuals without musical training. The perception of music involves complex brain functions underlying acoustic analysis, auditory memory and auditory scene analysis and processing of musical syntax. Moreover, music perception potentially affects emotion, influences autonomic nervous system, the hormonal and immune systems and activates (pre) motor representations.

Numerous studies have demonstrated superior performance of musicians in auditory processing as compared to non-musicians. Many studies are documented in literature showing that musical training improves basic auditory perceptual skills in turn resulting in behavioral enhancement (Jeon & Fricke, 1997; Koelsch et al., 1999; Oxenham et al., 2003; Tervaniemi et al., 2005; Micheyl et al., 2006; Rammsayer & Altenmuller, 2006) and neurophysiological responses (Brattico et al., 2001; Pantev et al., 2001; Schneider et al., 2002; Shahin et al., 2003 & 2007; Tervaniemi et al., 2005; Kuriki et al., 2006; Kraus et al., 2009). Similar and analogous to speech perception in noise musicians have life long experience in detecting melodies from background harmonies. For the acoustic stimulus in the presence of noise, musicians have demonstrated a more robust sub- cortical representation (Kraus et al., 2009). Musical practice not only enhances the processing of music related sounds but also influences processing of other domains such as language (Marques et al., 2007; Moreno et al., 2009; Parbery-Clark et al., 2009a; Schon et al., 2004).

In musical training discrimination of pitch intonation, onset, offset and duration aspects of sound timing as well as the integration of multisensory cues to perceive and produce notes are taught. Musical training facilitates musicians to pay more attention to the acoustic details of the stimulus than non-musicians (Musacchia et al., 2007). Music is a complex auditory task and musicians spend years fine tuning their skills. Previous research has documented neuroplasticity to musical sounds as a function of musical experience (Fujioka, Trainor, Ross, Kakigi, & Pantev, 2005; Koelsch, Schroger, & Tervaniemi, 1999; Musacchia, Sams, Skoe, & Kraus, 2007; Pantev et al., 1998; Pantev, Roberts, Schulz, Engelien & Ross, 2001).

As reported by previous studies music training benefits auditory processing not only in the musical domain, but also in the processing of speech stimuli (Musacchia et al., 2007; Schon, Magne, & Besson, 2004; Wong, Skoe, Russo, Dees, & Kraus,2007). Consistent reports across a range of studies that use different methods indicate that music training improves a variety of verbal and non verbal skills ranging from behavior to neurophysiology. These include working memory (Chan, Ho, & Cheung, 1998; Foregeard, Winner, Norton & Schlaug, 2008), processing of prosody and linguistic features in speech (Chandrasekaran, Krishanan & Gandour, 2009; Wong et al., 2007), phonological skills (Foregeard, Schlaug, et al, 2008), processing emotion in speech (Strait, Kraus, Skoe, & Ashley, 2009), auditory attention (Strait, Kraus, Parbery- Clark, & Ashley, 2010) and auditory Stream segregation (Beauvois, & Meddis, 1997).

Auditory temporal processing can be defined as the perception of sound or of the alteration of sound within a restricted or defined time domain.  It is the perception of short interval of time that each individual can discriminate between two auditory signals which is of about 2-3 ms. The studies suggest that exposure to sound during the first two years of life is important for the maturation of the structures of the central nervous system. Functional reorganization of the cerebral cortex is induced by musical training diary, used by professional musicians, as evidenced by many studies. Therefore, the contact with the music before the age of seven could contribute to the development of the primary auditory cortex more precisely the planum temporale. Also investigations of magneto encephalography on musicians have shown an increase in the left temporal plane. Thus musicians have better neural activation due to long term musical training.

Some studies have documented results on the role of nature vs nurture in the musicians. As evidenced by many previous studies, musical training if begun before the age of nine, in such individuals better development of planum temporale is seen thus emphasizing the importance of nurture in acquiring the musical skills. However, other studies argue that musical ability is innate and that musical training is not responsible for the improvement of planum temporale. However, it is confirmed that there is improvement of the planum temporale in relation to individuals who were exposed to early musical stimulus (Pantev et al. 2001). Another study by Ishll et al. (2006) shows that music has a positive influence on the development of the planum temporale, because according to their study, subjects were exposed to musical training (singing) over four years compared to amateur musicians without professional guidance, performed better on temporal resolution through the test Random Gap Detection Threshold (RGDT).

Speech perception in noise (SIN) is a complex task which requires the segregation of the target signal from the competing background noise. This is a complex task which is influenced with the noise particularly disrupting the perception of the fast spectro-temporal changes making perception of the acoustic signal difficult (Brandt & Rosen, 1980). Children with language-based learning disabilities (Bradlow et al., 2003; Ziegler et al., 2005) and hearing impaired adults (Gordon- Salant & Fitzgibbons., 2005) are especially susceptible to the negative effects of background noise. However, musicians are less affected and demonstrate better performance for SIN when compared to non-musicians (Parbery- Clark et al., 2009). This is because musicians exhibit enhanced subcortical encoding of sounds with both faster responses and greater frequency encoding when compared with non-musicians. The underlying neural representation of sounds is selectively strengthened by the musical experience reflecting the interaction between cognitive and sensory factors (Kraus et al., 2009), with musicians demonstrating better encoding of complex stimuli (Wong et al., 2007) as well as behaviorally relevant acoustic features (Lee et al., 2009).

The target acoustic signal is extracted by our auditory system by means of two processes. The first process involves partition of the acoustic input into separate auditory units. Second involves a mechanism for appropriately organizing these acoustic units over time. This internal process of segregating and subsequent grouping of an auditory stream is termed as auditory scene analysis (Bregman, 1990). Koffka in 1935 reported that the auditory scene analysis is based on the notion that pre-attentive processes use the Gestalt laws of organization - physical similarity, temporal proximity, and good continuity- to group sounds. Sounds with similar frequency and spatial location are more likely to be grouped together as auditory units. Indeed the perception of speech perception in noise depend on both frequency and spatial location cues. The ability to hear two streams or perceptual streaming is facilitated when concurrently presented complex tones are separated by as little as one semitone. For example, when the fundamental frequencies of the simultaneously presented vowels differ, the performance improved (Scheffers., 1983; Assmann & Summefield., 1990). This can help us to understand why speech recognition in noise is more difficult when the target and the background speakers are of the same sex, and the fundamental frequencies of different voices are consequently closer in frequency. Even small frequency differences between speakers’ voice can be used as cues to aid speaker differentiation (Treisman., 1964; Brkox & Nooteboom, 1982).

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

Musicians are trained in most of the time attending and manipulating multiple auditory streams. In addition to processing simultaneously occurring melodies, musicians must also analyze the vertical relationships between harmonies. In addition to auditory scene analysis musicians also hone their abilities to conceive, plan, and perform music in real time. Literature has documented how musical training improves basic auditory perceptual skills thereby enhancing behavioral output (Jeon & Fricke, 1997; Koelsch et al, 1999; Oxenham et al. 2003; Tervaniemi et al. 2005; Rammsayer & Altemuller 2006) and neurophysiological responses (Brattico et al. 2001; Pantev et al. 2001; Schneider et al. 2002; Tervaniemi et al. 2005; Kraus et al. 2009). The musicians are able to use these perceptual benefits to facilitate concurrent sound segregation (Zendel & Alain 2009). Musical training does not limit the benefits specifically to musical perception but also transfers these enhancements to cross over to other domains, particularly language, suggesting that there exist shared neural resources for language and music processing (Patel 2003, 2007; Kraus & Banai 2007; Kolesch et al. 2008). For example, lifelong musical experience is linked to improved subcortical and cortical representations of acoustic features which is important for speech encoding and vocal communication (Magne et al. 2006; Schon et al. 2004; Marques et al. 2007; Musacchia et al. 2007, 2008; Chandrashekaran et al. 2008; Strait et al. 2009). Also, it has been reported that musical experience improve verbal ability (Foregeard et al. 2008), verbal working memory and verbal recall (Chan et al. 1998; Brandler & Rammsayer 2003; Ho et al. 2003; Jackobsen et al. 2003). As a result of the musician’s extensive experience with auditory stream analysis within the context of music, more honed auditory perceptual skills as well as greater working memory capacity, musicians seem well equipped to cope with the demands of adverse listening situations such as speech in noise.

The studies mentioned above have highlighted enhanced auditory perceptual skills in trained musicians when compared to non-musicians. But there are only very few studies which have studied the effect of both music training and familial background of music on temporal resolution abilities and speech in noise among. Hence, in this regard the present study was aimed to study the same in musicians with and without familial background of music.

**METHOD**

**Subjects**

Two groups of subjects were selected, 20 children in the age range of 12-16 years who are undergoing formal musical training formed group ‘A’ and 20 age matched children who are not trained for music formed group ‘B’. All the subjects participated had normal air conduction and bone conduction thresholds (≤ 15 dB HL) at all octave frequencies from 250 Hz to 8000 Hz. An informal Questionnaire was administered to all participants in order to obtain the information like experience in the musical field. All testing was carried out in a sound treated double room situation as per the standards of ANSI S3.1 (1991).

**Instrumentation**

The following instruments were used in the present study:

1. Orbiter 922 (Madsen Electronics, Denmark), two channel audiometer, calibrated as per ISO 389, with supra-aural headphones (Telephonics TDH-39) housed with MX-41/AR ear cushions with audio cups and a bone vibrator (Radio ear B71) were used to assess the pure tone threshold.
2. GSI Tympstar (Grason-Stadler Inc. USA) middle ear analyser was used for tympanometry and reflexometry.
3. A laptop (Acer Aspire 5050) was used to deliver the stimulus for Gap Detection Test (GDT), Temporal Modulation Transfer Function (TMTF) and Quick Speech in Noise (QuickSIN) - Kannada.

**Stimuli:**

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

**Procedure:**

*Temporal Modulation Transfer Function (TMTF)*

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

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

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

*Gap Detection Threshold (GDT)*

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

*Quick Speech Perception in Noise - Kannada*

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

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

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

To examine the relationship between numerical variables obtained from TMTF, GDT and quick SIN Kannada, Pearson correlation was administered. The level of significance for statistical test was 5%, i.e., p<0.05.

**Results**

To obtain the relationship between temporal resolution tests and quick speech in noise test, Pearson correlation was administered. The data were appropriately tabulated and statistically analyzed using SPSS (Version 18) software. Table 1 and 2 shows the relationship of group ‘A’ and ‘B’ respectively for TMTFs, GDT with Quick SPIN test.

Table 1. Pearson correlates analysis across TMTF, GDT and Quick SPIN for group ‘A’

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Group ‘A’ | 4Hz | 8Hz | 16Hz | 32Hz | 64Hz | 128Hz | GDT |
| QuickSPIN  Correlation Coefficient (ρ) | 0.404\*  0.010 | 0.255  0.112 | 0.455\*  0.003 | 0.410\*  0.009 | 0.671\*  0.010 | 0.582\*  0.010 | 0.162  0.318 |

\*Correlation (ρ) is significant at the 0.05 level

Table 2 Pearson correlates analysis across TMTF, GDT and Quick SPIN for group ‘B’

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Group ‘B’ | 4Hz | 8Hz | 16Hz | 32Hz | 64Hz | 128Hz | GDT |
| QuickSPIN  Correlation Coefficient (ρ) | 0.351\*  0.026 | 0.309  0.052 | 0.441\*  0.004 | 0.405\*  0.010 | 0.650\*  0.010 | 0.547\*  0.010 | 0.101  0.535 |

\*Correlation (ρ) is significant at the 0.05 level

**DISCUSSION**

The present article attempts to establish the relationship between the temporal resolution tests and speech in noise tests among musicians versus individuals with no musical training. From the results, it is observed that both the groups ‘A’ and “B’ showed statistically significant positive correlation between test of temporal resolution and speech in noise.

Group ‘A’ indicates stronger positive relation between temporal resolution tests and speech in noise tests than compared to that of group ‘B’. The correlations are significant at the 0.05 level for both the groups except for 8Hz modulation rate and GDT scores. In the absence of the specific literature on TMTF in musicians it is difficult to explain the result found in the present study which has shown no difference at 8Hz rate. If tests to tap auditory working memory were included in the study those would have thrown light on providing explanations to results. Additionally it would have contributed to know the underlying processes involved. But in general, according to Ishll, et al. (2006), when random gap detection test was administered on musicians and non-musicians, the gap detection thresholds were better in trained musicians when compared to non-musicians. This helps to draw the conclusion that the temporal resolution abilities are better in musicians when compared to non-musicians.

According to Ohinshi et al (2001), music training can induce functional reorganization of the cerebral cortex. Therefore, musical training initiated early in life, before the age of seven contributes to the development of primary auditory cortex and more precisely the planum temporale. When the GDT was compared between the two ears within the group there was a statistical significant difference at 5 % level of significance. Hence, the present study warrants exploring in depth and elucidating the reason for such relationship between the temporal resolution processes and speech in noise.

**REFERENCES**

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

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

Assmann, P. F., & Summerfield, Q. (1990). Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies. *Journal of Acoustical Society of America, 88*, 680-697.

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

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

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

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

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

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

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

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.

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

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

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

Gordon-Salant, S., Fitzgibbons, P. J. (2005). Recognition of multiply degraded speech by young and elderly listeners. *Journal of Speech Language and Hearing Research, 38*, 1150-1156.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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.

Ohinshi, T., Matsuda, H., Asada, T., Aruga, M., Hirakata, M., Nishikawa, M. (2001). *Cerebral Cortex, 11(8)*, 75-760.

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

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

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

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

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

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

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

Schlaug, G., Norton, A., Overy, K., & Winner, E. (2005). Effects of music training on the child’s brain and cognitive development. *Annals of New York Academy of Sciences, 930*, 281-299.

Schneider, P., Scherg, M., Dosch, H. G., et al. (2002). Morphology of Heschl’s gyrus reflects enhanced activation in the auditory cortex of musicians. *Natural Neuroscience, 5*, 688-694.

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

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

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

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

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

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

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

Wong, P. C., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English speaking adults. *Applied Psycholinguistics, 28*, 565-585.

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

Zendel, B. R., Alain, C. (2009). Concurrent sound segregation is enhanced in musicians. *Journal of Cognitive Neuroscience, 21*, 1488-1498.

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

**Acknowledgements**

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