PERCEIVING EMOTION IN EXPRESSIVE PIANO PERFORMANCE: A FUNCTIONAL MRI STUDY

Dinesh G. Nair[1]*, Edward W. Large*[1]*, Fred Steinberg*[1,2] *and J A Scott Kelso*[1]

[1]Center for Complex Systems and Brain Sciences, Florida Atlantic University, Boca Raton, Florida, USA.

[2] University MRI of Boca Raton, Boca Raton, Florida, USA.

ABSTRACT

We aimed to identify brain areas involved in responding to affect communicated by expressive piano performance. Our subjects listened to two versions of Chopin's Etude in E major, Opus 10, No. 3. The first version was an expressive performance, recorded by a highly trained musician on a computer-monitored piano. Our control was a computer-generated, mechanical performance of the same composition. Data analysis revealed differential brain activation in the two listening conditions. The expressive performance elicited greater activation in anterior cingulate, right temporal pole, right inferior frontal gyri, inferior parietal lobe and superior temporal gyri, areas that have been associated with emotion, attention, speech perception. The mechanical performance elicited greater activation in cerebellum, parahippocampal gyrus, supplementary motor area and dorsolateral prefrontal cortex, areas primarily involved in motor and sequencing tasks. Our results confirm that expressive music performance communicates affect beyond the melody, harmony, tonality, and rhythm of the notated composition. Our observations also suggest that the perception of emotion in music shares neural resources with the perception of emotion in speech, and that these pathways may be different from those recruited during other types of emotional experience.

1. INTRODUCTION

Listening to music involves perceiving sound stimuli, grouping them into patterns and relating these patterns to one another. But understanding musical expression calls for processes by which the listener experiences sound patterns as feelings and emotions. Music is believed to evoke a wide range of affective states in the absence of external associations [15]. Listeners, whether musically trained or not, are in general able to name the emotion that a musical excerpt was intended to convey, even across cultures [1]. However, the study of emotions conveyed by musical excerpts is susceptible to a number of problems. First, because music flows through time, it is difficult to pinpoint musical processes that evoke particular affective responses. Second, although some music is meant to convey specific emotions, a great deal of music is not intended to convey stereotypical emotions at all [11]. To paraphrase Kraut, listeners rarely experience envy, indignation, love, or fear when listening to uptempo Ornette Coleman performances [6]. Nevertheless, such musical experiences can arouse intense affective responses. Music is also known to evoke physiological responses and certain structural properties of music have even been linked to specific physiological responses. For instance, Sloboda [15] showed that tears were evoked by melodic appoggiaturas and relatively sudden changes in harmonies evoked shivers. Krumhansl [7] observed that sad music resulted in increased systolic, diastolic and mean arterial pressures and decreased heart rate, skin conductance and finger temperature. But these physiological responses tell us neither about the nature of the processes underlying emotional experience nor their relation to the piece of music.

The neural correlates of musical processing have been studied widely in the last decade and researchers have identified brain areas involved in the detection of pitch, contour, rhythm, meter and other structural aspects of music [10, 17]. Fewer studies have investigated neural correlates of emotional responses to music. Peretz et. al., [14] studied a patient with amusia (but without aphasia) who exhibited normal emotional judgment for a piece of music but had gravely impaired music processing abilities, and suggested the existence of separate neural pathways for emotional interpretation compared to structural interpretation of music. Using Positron Emission Tomography (PET) Blood et. al., [2] showed that when subjects listened to musical passages which varied systematically in the degree of dissonance, cerebral blood flow changes in the right parahippocampal gyrus and precuneus regions correlated with increasing dissonance while activity in the orbitofrontal, subcallosal cingulate and the frontal polar cortex correlated with decreasing dissonance. These were distinct from the areas involved in the analysis of structural components of music.

Performers use various cues to convey emotion and meaning to listeners, collectively these are called performance expression. In piano performance, the cues are limited mainly to fast time-scale fluctuations in timing, (rubato and articulation) and intensity (dynamics). Similar fluctuations are also observed in speech communication and they are known to communicate many types of information to listeners. The ways in which performance timing and intensity variations communicate musical structure (e.g. phrasing, meter) has been extensively studied [12] and has even been modeled in some detail [9]. Yet we know few details about how listeners perceive affect in music performance.

In this functional Magnetic Resonance Imaging (fMRI) study, participants listened to two versions of the same piece, one performed by a highly trained musician, and the other generated by computer to conform as precisely as possible to the notated composition. Thus, the two listening conditions were matched for melody, harmony, tonality, and rhythm. They differed only along performance parameters used by pianists to communicate with listeners: dynamics, articulation and rubato. We address the following questions. Are different brain areas activated when listening to expressive versus mechanical performances?

Can we draw inferences from our observations about how musical performance conveys emotion and meaning? Does communication of affect in music involve the same brain areas as other types of emotional responses, or is musical communication special in some way?

2. METHOD

2.1 Participants

Our listeners were four musicians (mean performance experience of 31.5 years; range: 25-40 yrs). Informed consent was obtained from all subjects after explaining to them the nature of the experiment. Subjects filled out a questionnaire after the experiment in which they responded to questions about their musical experience and their familiarity with the piece of music.

2.2 Stimuli

We used Chopin's Etude in E major, Opus 10, No. 3 as our stimulus. This piece was performed by a senior piano major on a Kawai CA 950 digital piano, and recorded into Studio Vision running on a Macintosh G3 computer (Mac OS 9.0.4). To conform to the block design of the functional MRI paradigm, the performance was divided into six 30-45 second listening blocks. Blocks were chosen to conform to musical sections or subsections to cause minimal interruption to the natural flow of the music. Next, a mechanical performance was synthesized on the computer by changing the onset time and duration of each note to precisely match that of the musical notation, the onset velocity (MIDI) of each note was set to 64, and pedal information was eliminated. The mechanical version was then divided into listening blocks, and each block was matched for mean tempo with the corresponding block of the expressive performance by time stretching or compression. Listening blocks were interspersed with 30-second blocks of silence. The stimuli were played back via MIDI, through the Kawai CA 950, and recorded on a Sony PCM 2500B digital tape recorder.

2.3 Equipment

Whole brain fMRI data acquisition was carried out using a 1.5 Tesla Signa scanner (General Electric Medical Systems, Milwaukee, USA). The stimulus was played to the subjects from the digital tape through non-magnetic tubes and headphones (Avotec Inc). Headphones were custom-modified to deliver sound directly into the external auditory canal, by attaching soft-tipped earplugs of a Littmann™ Cardiology III stethoscope to a thick plastic tube that was shaped to precisely match the shape of the stethoscope, and then inserted into the sound-protective Avotec headphone shells. Sound barriers (Sonex™, 1 inch thick, mean 30 dB attenuation within the frequency range of our performance) were used to insulate the auditory junction box and the head coil of the magnet from scanner noise.

2.4 Procedure

Each condition (performed and notated) lasted for 6 minutes and 33 seconds, and was comprised of 6 periods of activation (ON, listen, 30-45 sec) during which subjects listened to music and 6 baseline (OFF, rest, 30 sec) periods in which subjects heard only the ambient machine noise. Subjects were instructed to close their eyes and carefully listen to the performance. The subject's head was supported by a comfortable foam mold and head movement was further minimized using foam padding and forehead restraining straps. Scanning started with the acquisition of full head, 3D SPGR (spoiled gradient) anatomical images, with the following imaging parameters: Field of view (FOV) of 26 cm, frequency-phase matrix size = 256 x 256, repetition time (TR) = 34ms, echo time (TE) = 5 ms, flip angle (FA) 45^o, slice thickness 2mm, and one excitation (NEX) per phase encoding step. For each subject, T2*-weighted gradient echo, echo planar multi-slice datasets were acquired during ON and OFF periods, with a TR of 3 sec, TE 60 ms and FA = 90^o (20 axial slices, matrix = 64 x 64, FOV = 24 cm, slice thickness = 5 mm and inter-slice gap = 2.5 mm). Thus the voxel size was 3.75 x 3.75 x 7.5 mm. High-resolution background images (same 20 slices, matrix = 256 x 256, NEX = 2) were also acquired to overlay the functional data. During the structural scans, subjects listened to a different piece of music recorded using the same settings, so that a comfortable loudness level for the stimulus was achieved. This also accustomed them to the process of listening to music in an MRI environment.

2.5 Data Analysis

The software used for analysis was AFNI (Analysis of Functional NeuroImages, Medical College of Wisconsin), [4]. We first performed movement correction of the functional datasets by using the Fourier method in AFNI. The raw time series were low pass filtered (cut off = 0.07 Hz) and spatially filtered using a Gaussian kernel (FWHM = 6mm) to enhance signal to noise ratio. Alternating periods of baseline and listening-related activation were modeled using boxcar reference waves shifted by 3, 6 or 9 seconds respectively to account for the hemodynamic response delay. This delay was determined by examining the raw time series data. Regions of task-related activity were determined by cross correlation of the image time series with the reference waveforms. The first stage of analysis used a thresholding procedure in which voxels with correlation coefficients greater than or equal to a threshold of 0.5 were identified and retained for further analysis. The correlation values were then converted to z-scores for all task conditions and all subjects. The resulting data were transformed into the Talairach and Tournoux stereotaxic space [16] for comparison across subjects. The mean intensity of activation across all subjects in the expressive performance was compared with that in the mechanical performance, to look for differences in brain activation across the two tasks. The significance of these differences was determined by using a paired t-test ($p<0.05$). In order to correct for multiple comparisons, we used probability thresholding in combination with cluster size thresholding. Only those voxels above $p<0.05$ (corresponding to $t = 3.16$) within a radial distance of 2mm from an active voxel and those that formed a volume of at least 1050 µl (10 times the volume of one original voxel) were labeled as an active cluster. Images were created by mapping voxel t-values to colors using a scale from red (minimum) to yellow (maximum) when expressive > mechanical and blue (minimum) to cyan (maximum) when mechanical > expressive.

3. RESULTS AND DISCUSSION

The results of the paired t-test between the two conditions (expressive and mechanical) are shown in Figures 1 and 2. When listening to the mechanical performance, subjects showed stronger activation of right parahippocampal gyrus (PHG; Fig. 1a), and ventral posterior cingulate gyrus (not shown) during the mechanical performance. A previous PET study showed that blood flow to PHG was positively correlated with increasingly unpleasant musical stimuli [2]. Indeed, our subjects reported finding the mechanical performance somewhat unpleasant to listen to. Further, ventral posterior cingulate has been implicated in the processing of emotion, and its activity is known to correlate with increasingly painful stimuli. It may also be recalled here that extensive connections exist between the parahippocampal gyrus and the cingulate cortex, both being components of the limbic system and involved in the Papez circuit [5].

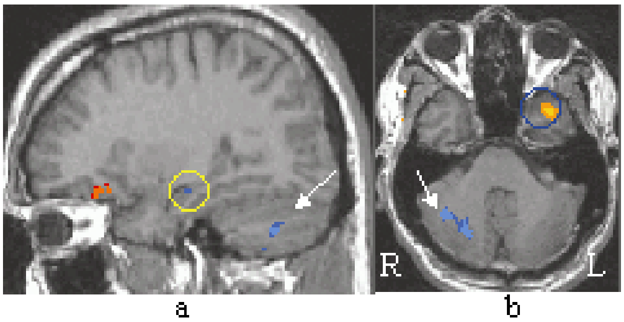

Figure 1. Increased intensity of activation was seen in the right parahippocampal gyrus (yellow oval) and the right cerebellum (white arrows) while listening to the mechanical performance. The red voxels in the left panel represent inferior frontal gyrus (BA 11, BA 47 or frontal operculum) and the yellow voxels in the right panel represent left temporal pole (BA 38), both of which showed higher intensity of activation when subjects listened to the expressive performance. R - right; L – left.

Increased intensity of activation was also observed in dorsolateral prefrontal areas (Brodmann’s area [BA] 9, Fig. 2b), cerebellum (Fig. 1, arrows), and SMA (not shown) during the mechanical listening task. The mechanical performance had a strong, predictable rhythm, thus activation in these brain areas may have been due to some form of mental beat-following behavior. This could have recruited a loop involving the STG, limbic, and dorsolateral prefrontal areas, with projections to the cerebellum. Direct connections between these areas are known to exist in the human brain [5].

During the expressive listening condition, higher intensity of activation was observed bilaterally in both the transverse temporal gyri (BA 41 & 42) and the superior temporal gyri (BA 22; Fig. 2a), which included parts of primary, secondary, and associative auditory cortices. Earlier studies have provided evidence for the role of these areas in processing pitch, contour, rhythm and meter [10, 17]. Thus, this observation was somewhat surprising since these areas were expected to be equally active during both listening conditions. We return to this observation momentarily.

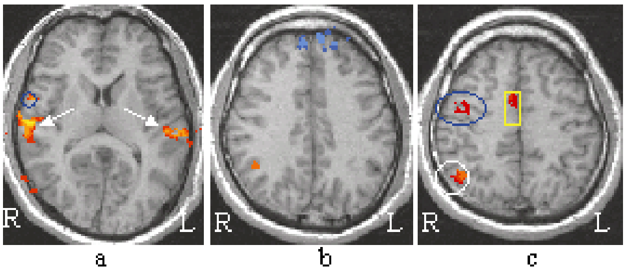

Figure 2. Comparison of brain activation (t-test; $p<0.05$) between the two tasks revealed increased activation in bilateral superior temporal gyri – BA 22 and bilateral transverse temporal gyri - BA 41/42 (white arrows, panel a), right middle temporal gyrus, right inferior frontal gyrus - BA 44 (blue oval panel a), right supramarginal gyrus and right angular gyrus– BA 40, 39 (red voxels, panel b), right cingulate gyrus - BA 24, 32 (rectangle, panel c) right precentral gyrus (blue oval, panel c) and right inferior parietal lobule (white circle, panel c). Bilateral superior and medial frontal gyri (BA 9; blue voxels, panel b) were more active when subjects listened to the mechanical version. Panels a, b and c correspond to slices at z coordinates 10, 34 and 46 (in the vertical axis) of the Talairach coordinate system. R – right; L – left.

Higher intensity of activation was seen in the right anterior cingulate cortex (ACC, BA 24 & 32, Fig. 2c) during the expressive performance. ACC has been implicated in a variety of functions including emotion, attention, novelty and error detection [3, 5]. Increased activation in ACC may reflect affective or emotional responses of listeners to the expressive performance. This would be consistent with our additional observation of increased activity in the temporal pole (BA 38, Fig. 1b), which also forms part of the limbic system and has functional connectivity with ACC [5]. Additionally, it is likely that expressive timing and intensity variations resulted in increased levels of attention, thus recruiting neurons in the ACC. This interpretation is consistent with our observation of increased activity in auditory areas, which may also play a role in affective processing. Finally, in the expressive listening condition, higher intensity activation was observed in the right inferior frontal gyrus (BA 44; Fig. 2a), right supramarginal gyrus (BA 40), right angular gyrus (BA 39; Fig. 2b), right inferior parietal lobule (Fig. 2c) and right frontal operculum (Fig. 1a). These areas play a variety of roles in speech and language processing. For example, it has been shown that patients with lesions in the right inferior parietal lobe fail to appreciate aspects of a verbal message that are conveyed by prosodic cues [5]. Activation of these areas while listening to expressive music performance implies sharing of neural resources that are important in linguistic function, including the processing of both prosody and semantics. This observation is also in agreement with previous neurological evidence showing that the processing of music-like sound patterns involves the same neural resources as the processing of prosodic patterns in speech [13].

4. CONCLUSION

We observed differential brain activation depending upon whether participants listened to a mechanical or expressive performance of the same musical composition. The simple rhythm of the mechanical performance preferentially activated regions involved in timing and movement planning. The microstructure of the expressive performance, on the other hand, recruited an intricate neural network that functionally links bilateral auditory and auditory association areas with limbic/paralimbic (cingulate, parahippocampal gyrus, temporal pole) and speech processing areas (inferior frontal, frontal operculum, inferior parietal lobe). Interestingly, although we observed activation of certain emotion areas such as the cingulate, other important limbic areas such as the amygdala did not show significant activity. Although preliminary, this finding raises the possibility that music communicates affect in a way that is distinct from many other emotional experiences. It would also offer indirect support for the theory that musical experiences tend to produce non-specific affective arousal, which may or may not be interpreted as emotion, rather than directly communicating specific identifiable emotions [11].

Meyer's approach holds that violation of expectancy forms the basis for the communication of emotion and meaning in music [11]. If rhythmic expectancies are violated by expressive timing deviations [8, 9] this would provide a theoretical basis for increased emotional response to the expressive performance. It would also explain the apparent increase in attention, since violation of temporal expectancy would result in attentional capture [8]. In addition, the extensive activation of neural areas previously associated with prosody suggests that similar processes are at work in communicating affect in speech. An expressive music performance is more than a sonic realization of a musical score. It is widely understood that expressive performance communicates aspects of musical structure [9, 12]. In this study we observed, for the first time, recruitment of limbic/paralimbic areas in response to music performance, implying communication of affect. We also found recruitment of neural structures related to components of attention and speech processing, which may help us to better understand this process of musical communication between performer and listener.

5. ACKNOWLEDGEMENTS

This research was supported by NSF grant BCS-0094229 awarded to EWL and NIMH grant MH 42900.

6. REFERENCES

1. Balkwill, L-L., Thompson, W.F., "A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues", *Music Percept*., Vol 17(1): 43-64, Fall 1999.
2. Blood, A.J., Zatorre, R.J., Bermudez, P. and Evans, A.C. "Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions", *Nat. Neurosci*., Vol 2 (4): 382-387, April 1999.
3. Carter, C.S., Braver, T.S., Barch, D.M., Botvinick, M.M., Noll, D. and Cohen, J.D. "Anterior Cingulate cortex, Error detection and online monitoring of performance", *Science*, Vol 280: 747-749, May 1998.
4. Cox, R.W., "AFNI: software for analysis and visualization of functional magnetic resonance neuroimages", *Comput. Biomed. Res*., Vol 29: 162-173, 1996.
5. Kandel, E.R., Schwartz, J.H. and Jessell, T.M., "*Principles of Neural science*", Appleton and Lange, Norwalk, Connecticut, 1991.
6. Kraut, R. "On the possibility of a determinate semantics for music", In M. R. S. Holleran (Eds.) *Cognitive Bases of Musical Communication*, Washington, DC: APA, 1992.
7. Krumhansl, C.L., "An exploratory study of musical emotions and psychophysiology", *Can. J. Exp. Psych.*, Vol 51(4): 336-352, 1997.
8. Large, E.W., and Jones, M. R. "The dynamics of attending: How we track time-varying events", *Psych. Rev*., Vol 106 (1): 119-159, 1999.
9. Large, E.W., and Palmer, C., "Perceiving temporal regularity in music", *Cogn. Sci*., Vol 26: 1-37, 2002.
10. Liegeois-Chauvel, C., Peretz, I., Babai, M., Laguitton, V. and Chauvel, P., "Contribution of different cortical areas in the temporal lobe to music processing", *Brain*, Vol 121, 1853-1867, 1998.
11. Meyer, L.B., "*Emotion and meaning in music*", Chicago: University of Chicago Press, 1956.
12. Palmer, C., "Music performance", *Ann. Rev. Psych*., Vol 48, 115-138, 1997.
13. Patel, A., Peretz, I., Tramo, M., and Labrecque, R. "Processing prosodic and musical patterns: A neuropsychological investigation", *Brain Lang*., Vol 61: 123-144, 1998.
14. Peretz, I., Gagnon, L. and Bouchard, B; "Music and emotion: perceptual determinants, immediacy and isolation after brain damage", *Cognition*, Vol 68: 111-141, 1998.
15. Sloboda, J.A., "Music structure and emotional response: some empirical findings", *Psychol. Music*, Vol 19: 110-120, 1991.
16. Talairach, J., Tournoux, P., "*Co-planar Stereotaxic atlas of the brain*", New York: Thieme, 1988.
17. Zatorre, R.J., Evans, A.C., and Meyer, E., "Neural mechanisms underlying melodic perception and memory for pitch", *J Neurosci*,. Vol 14 (4): 1908-1919, 1994.