Cortical Activation in Superior Temporal Gyrus and Fusiform Gyrus Modulated by Congruence of Emotional Content in Music and Face

J-W. Jeong[1], V. Diwadkar[2], C. D. Chugani[3], H. T. Chugani[4], and D. C. Chugani[5]

[1]Pediatrics, Neurology, Wayne State University, Detroit, Michigan, United States, [2]Psychiatry, Behavioral Neuroscience, Wayne State University, Detroit, Michigan, United States, [3]Florida Gulf Coast University, [4]Pediatrics, Neurology, Radiology, Wayne State University, Detroit, Michigan, United States, [5]Pediatrics, Radiology, Wayne State University, Detroit, Michigan, United States

Introduction Emotional cues from the environment are obtained through multiple sensory modalities, and the integration of these emotionally laden sensory signals might enhance or diminish the emotional experience conveyed by each sense alone [1]. For example, the pairing of pictures and music conveying the same emotion appears to amplify the affective experience of the viewer. The neural basis of how affective cues from multiple sensory modalities are integrated in the brain is not well understood. It is plausible that emotionally consistent cues enhance or modulate the response of regions specific to each of the sensory domains (e.g., visual or auditory) in which the information is delivered, whereas affectively inconsistent cues may serve to interfere with the response of these areas. This study hypothesizes that neural mechanisms involved in listening to music may differ when presented together with visual stimuli that conveyed the same emotion as the music when compared to visual stimuli with incongruent emotional content. Based on above hypothesis, we present a novel functinal MRI paradigm combining happy and sad music with happy and sad faces [2] to examine the response of superior temporal gyrus (STG) and fusiform gyrus (FG) to assess the neural bases of integrating visual and auditory stimuli congruent and incongruent with regard to emotional content. In particular, we are interested in assessing whether congruence for emotion influenced the response of STG and FG as the heteromodal sites that integrate auditory and visual information.

Meterials and Methods Fifteen healthy, right-handed volunteers (age 22.8±3.4 years, 10 females) participated in the functional imaging experiment. All subjects reported having normal hearing, no neurological disorders or head injuries. The experiment was conducted in the context of a conventional block-design experiment. A block consisted of three emotional ON stimuli, music alone [30 second excerpt of *Blue Danube Waltz* by Strauss for happy music (HM), or 30 second excerpt of *Adagio for Strings* by Barber for sad music (SM)], face alone [30 second presentation of 10 happy face images at 3 second intervals for happy face (HF), or 30 second presentation of 10 sad face images at 3 second intervals for sad face (SF)], and music combined with faces where the music excerpt was played while presenting either congruent emotional faces ["HM+HF", "SM+SF"] or incongruent emotional faces ["HM+SF", "SM+HF"]. These ON stimuli were ordered randomly to minimize habituation and order effects. A fixation cross was presented for the OFF condition between every emotion ON stimulus. fMRI experiments were conducted on a 3-T GE Signa MR equipped with an 8-channel ASSET head coil. A total of 336 volumes were acquired using T2*-weighted echo-planer imaging sequence at TR = 3000ms, TE = 35ms, flip angle = 90°, field of view = 240 mm, matrix size = 64x64, slice thickness = 4 mm, and zero gap covering whole brain. The functional time series was realigned, normalized into a standard stereotaxic space using the MNI EPI template, and smoothed using an 8-mm FWHM Gaussian kernel filter to permit application of Gaussian random field theory and to facilitate inter-subject averaging. The eight experimental conditions, (HM, HF, "HM+HF", "HM+SF", SM, SF, "SM+SF", "SM+HF") were modeled with a box-car function convolved with a hemodynamic function using the SPM general linear model. The first-level analysis identified activation in individual subjects corresponding to the following main contrasts of interest: a) main effect of congruence for emotion: (HM+HF)+(SM+SF) > (HM+SF)+(SM+HF), b) main effect of incongruence for emotion: (HM+SF)+(SM+HF) > (HM+HF)+(SM+SF). Each contrast image was combined with the second level analysis using a one-sample t-test to determine group averaged activations. Inference was carried out based on family wise error (FWE) p-values that were adjusted for small volume correction implemented in SPM8. An *a priori* ROI analysis was applied to reveal significant effects of emotion congruence vs. incongruence in bilateral STG and FG. Subsequent to the imaging study, 11 of the 15 subjects underwent behavioral testing to assess their ratings for emotion of the music and face stimuli. For music trials, each musical excerpt was presented for 30 seconds through earphones. The subjects were instructed to rate the emotion on a scale of -7 (saddest) to +7 (happiest) after listening to entire musical excerpt. For face trials, each face was displayed for 3 seconds on a computer screen. Subjects were instructed to give the first value that came to mind using the same emotion scale. Faces were presented either alone or together with each of the musical excerpts. In order to determine whether presentation of music together with faces had an effect of rating of emotion in faces, ratings for faces alone were subtracted from ratings of faces under each music condition for each subject as follows: (HM+HF)-HF, (SM+HF)-HF, (SM+SF)-SF, (HM+SF)-SF. All p-values were estimated from a Student' paired t-Test with a two-tailed distribution.

Results and Discussion Significant activations were found in the STG and FG. These activations were differentially modulated by music and faces depending on the congruence of emotional content ($p_{FWE} < 0.05$ after small volume correction). There was a greater BOLD response in STG when the emotion signaled by the music and faces was congruent (Fig. 2, p=0.008/0.07 for left/right STG). Furthermore, the magnitude of these changes differed for happy congruence and sad congruence, i.e., the activation of STG when happy music was presented with happy faces was greater than the activation seen when sad music was presented with sad faces. In contrast, incongruent stimuli diminished the BOLD response in STG and elicited greater signal change in bilateral FG (Fig. 2, p=0.002 for left/right FG). Behavioral testing supplemented these findings by showing that subject ratings of emotion in faces were influenced by the emotion conveyed by the music (Fig. 3). When presented with happy music, happy faces were rated as more happy (p=0.051) and sad faces were rated as less sad (p=0.030). When presented with sad music, happy faces were rated as less happy (p=0.008) and sad faces were rated as sadder (p=0.002). These findings are broadly consistent with the actual effects of cross-modal perceptual effects called the "emotional McGurk effect" [3]. Emotional congruence across modalities enhances activity in auditory regions while incongruence appears to impact the perception of visual affect, leading to increased activation in face processing regions such as the FG. We suggest that greater understanding of the neural bases of emotional congruence across modalities can shed light on basic mechanisms of affective perception and experience, and may lead to novel insights in the study of emotion regulation and music therapy. **Reference** [1] Campanella S., Belin P. 2007. Trends Cogn Sci 11:535-43. [2] Gur R.C. et al. 2002. NeuroImage 16:651-62. [3] de Gelder B., Bertelson P., 2003. Trends Cogn Sci. 7:460-7.

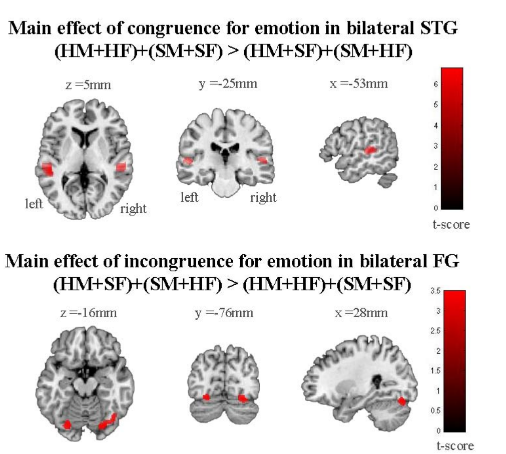

Fig.1: Main effect of congruent and incongruent stimuli.

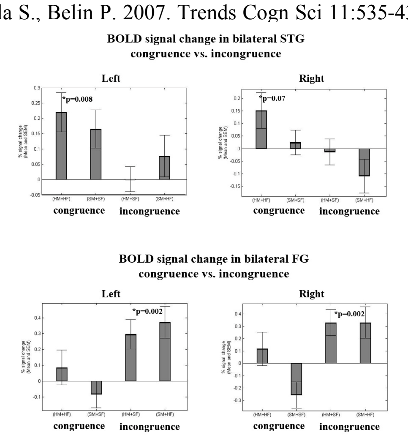

Fig. 2: BOLD signal changes in ROIs responding to congruent and incongruent stimuli. The percent signal change was evaluated at the peak of each cluster in Fig.1.

A

B

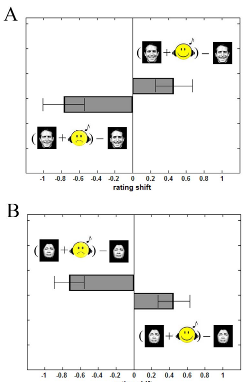

Fig. 3: Change in face emotion ratings by simultaneous presentation of music. **A**: Effect of happy and sad music on rating of happy faces. **B**: Effect of happy and sad music on rating of sad faces.