2012 International Conference on Biological and Life Sciences
IPCBEE vol.40 (2012) © (2012) IACSIT Press, Singapore

Brain Activities between Auditory and Visual Modalities in the Perception of Peace

Chia-Yen Yang[1,*] and Ching-Po Lin[2]
[1]Department of Biomedical Engineering, Ming-Chuan University, Taoyuan, Taiwan
[2]Institute of Neuroscience, National Yang-Ming University, Taipei, Taiwan

Abstract. Multisensory integration involves combining information from different senses to create a perception. The diverse characteristics of different sensory systems make it interesting to determine how cooperation and competition contribute to emotional experiences. Therefore, the aim of this study was to estimate the bias from the match attributes of the auditory and visual modalities and to depict specific brain activity frequencies related to the audiovisual modality in a peaceful mood by using magnetoencephalography. Our results revealed contrast in perceptual bias during multimodality processing of peaceful consciousness between auditorily and visually dominant subjects. Notably, hemispheric lateralization was also apparent in the alpha and beta oscillations, which might govern simple or pure information in the right brain but complex or mixed information in the left brain.

Keywords: Dominance, Audiovisual modality, Magnetoencephalography, Theta, Alpha, Beta

1. Introduction

Our environment is full of diverse stimuli (e.g., sound and light) that can be perceived by separate human physical senses to induce affective states. A successful social life requires the multitude of information from different sensory systems to be collected, analyzed, and interpreted to form unified percepts [1]. Recent investigations of multisensory integration suggest that responses to a stimulus in one modality (e.g., auditory, visual, or touch) can be influenced by the occurrence of a stimulus in another modality [2,3]. Integrating such sensory inputs is critically dependent on the congruency relation of interest, which may enhance or degrade the linkage of environmental events [4]. For example, when a subject views peaceful scenery accompanied by peaceful music, this is most likely to be perceived as peaceful, whereas when a subject views fearful scenery accompanied by peaceful music, it may be perceived as either fearful or peaceful. In other words, conflicting information from two modalities will be combined into an ambiguous affection.

Because of the divergence among different sensory systems, the weightings from each modality may not be equal in the fusion process [5]. Some studies have suggested that visual information is more important in the decoding of emotional meaning (e.g., [6]), whereas others have suggested that auditory information is more efficient for communicating emotional state (e.g. [2]). All these contradictions have demonstrated that the influences of the senses depend on both the stimulus properties and the task necessaries in everyday life [7,8]. Moreover, descriptions of the dominant modality probe the importance of intermodality discrepancies, which may vary with the situation [9]. Since auditory and visual sensory systems usually provide redundant information, the presence of cooperation and competition in the audiovisual modality is an interesting issue in emotional experiences.

Therefore, the aim of this study was to estimate the bias from the match attributes of the auditory and visual modalities and to depict specific brain activity frequencies (theta, alpha, beta, and gamma) related to the audiovisual modality in a peaceful mood by using magnetoencephalography (MEG).

2. Methods

2.1. Subjects

Eleven healthy males who were ophthalmologically and neurologically normal participated in this study (two of them who were later considered to be auditorily dominant subjects were not included in the same group for further statistical analyses). Their ages ranged from 22 to 42 years (28.0±6.0 years, mean±SD). The visual acuity of all subjects was corrected, where necessary, to the normal range. Informed consents that had been approved by the local Ethics Committee of Yang-Ming University were obtained from all the participants.

2.2. Stimuli and Stimulation Procedure

Before the recording session, 94 healthy males constituting another group of subjects were asked to rate their perceived feelings from 1 to 10 points of 10 (5 peaceful and 5 fearful) 90-s video clips and 10 (5 peaceful and 5 fearful) 90-s music clips. Clips of both emotional states that received scores of similar and at least 5 points were used in the subsequent MEG experiments. These stimuli formed a total of seven stimulus paradigms that consisted of three modalities (auditory, visual, and audiovisual) and two feelings (peaceful and fearful) as follows: (1) listening to peaceful (Ap) or fearful (Af) music without watching a film, (2) watching a peaceful (Vp) or fearful (Vf) film without listening to music, (3) watching a peaceful film while listening to peaceful music (ApVp), (4) watching a peaceful film while listening to fearful music (AfVp), and (5) watching a fearful film while listening to peaceful music (ApVf). The experimental procedure was separated into blocks by a break interval of about 1.5 min. During the entire experiment, the subject lay supine in a comfortable and stable position. A 1000-Hz pure tone was presented for 10 s (250 ms on/ 250 ms off) prior to listening to the music clip, while a checkerboard pattern was presented for 10 s (250 ms on/ 250 ms off) prior to watching the film clip. Thereafter, in block one, the paradigms (Ap), (Vp), and (ApVp) were applied sequentially, each followed by a 15-s response time for labeling the emotional expressions of sadness, happiness, peacefulness, fear, anger, or no emotion, and for scaling the score (ranging from 1 to 4 points). The participants were instructed to indicate their scores as accurately and quickly as possible after the paradigm had been presented. Similar procedures were followed in block two [i.e., rewind(Af)+rewind(Vp)+(AfVp) (data in rewind stimulus conditions were not analyzed in this study)] and block three [i.e., rewind(Ap)+rewind(Vf)+(ApVf)]. These visual and auditory stimuli were generated in MATLAB (The MathWorks, Natick, MA) using functions provided by the Psychophysics Toolbox [14,15] on a personal computer, and projected onto a mirror by a projector and delivered with binaurally inserted earphones via a silicone-tube system, respectively.

2.3. Data Acquisition

Visual evoked fields were recorded with a whole-head 160-channel coaxial gradiometer (PQ1160C, Yokogawa Electric, Tokyo, Japan). The magnetic responses were filtered by a bandpass filter from 0.1 to 200 Hz and digitized at a sampling rate of 1000 Hz. For off-line analysis, the nonperiodic low-frequency noise in the MEG raw data was reduced using a continuously adjusted least-squares method [16]. The resulting data were then filtered by a bandpass filter from 0.5 to 50 Hz after removing artifacts with amplitudes exceeding 3000 fT/cm in the MEG signals.

Anatomic image data was acquired from human subjects using a GE Healthcare Signa 1.5-T Excite scanner (General Electric, Milwaukee, WI) with an eight-channel head coil. The imaging parameters were as follows: TR = 8.548 ms, TE = 1.836 ms, TI = 400 ms, flip angle = 15°, field of view = 256 × 256 mm2, matrix size = 256 × 256, and no gaps. This gave an in-plane resolution of 1 mm2 and a slice thickness of 1.5 mm.

2.4. Data Analysis

The exported MEG data were preprocessed to remove eye movements, blink artifacts, and ECG activity by the FastICA algorithm [17]. The cleaned data were then segmented into epochs of 2 s beginning at the start of the segment (the first 10 s of the original data in each paradigm from all subjects were ignored to eliminate the possibility of attention transients associated with stimulus initiation). For each segment, two types of analysis were applied to the epoched data. The first produced an averaged spectrum by using wavelet transformation for each channel. The spectral power was calculated on a per-0.5-Hz basis and then classified into the following ten brain regions: 1, left frontal; 2, right frontal; 3, left frontotemporal; 4, right frontotemporal; 5, left temporal; 6, right temporal; 7, left central; 8, right central; 9, left occipital; and 10, right occipital. All these values were transformed into T-scores for each frequency band in order to compare different types of stimulations. Student's *t*-test was applied to analyze the score values for the different modalities for various brain waves (four frequency bands) and areas (ten regions). When the probability value (p) was <0.05, the difference was considered to be statistically significant.

3. Results

The responses measured in the MEG experiments revealed that all 11 subjects had the same feelings about the unimodal stimulus clips (i.e., feeling peaceful about Ap or Vp, and feeling fearful about Af or Vf), and rated the same scores for different modalities in the same states (e.g., 4 for Ap and Vp, but 3 for Af and Vf). However, nine of the subjects considered AfVp to be peaceful and ApVf to be fearful, while two of them made the opposite assignments. Figure 1 shows audiovisual maps of the theta, alpha, beta, and gamma activities in ten different regions from those nine subjects. Each dot represents one frequency for one subject. The distributions were concentrated around the diagonal (indicating equal weighting of the two modalities) but with a slight downward tendency, especially in the theta band around the frontal and left frontotemporal areas. Table 1 lists the mean ratios of unimodalities (i.e., Ap/Vp) and the mean differences between bimodalities [i.e., (ApVp – AfVp)/(ApVp – ApVf)]. The ratios were higher for the single modalities than for the combined audiovisual modality. ApVp differed significantly ($p<0.05$) from Ap in the theta and beta bands around the left occipital area, and from Vp in the theta band around the left frontotemporal and left occipital areas, in the beta band around the left frontotemporal and the occipital areas, and in the gamma band around the left frontotemporal and right occipital areas. Significant differences ($p<0.05$) were also found between ApVp and ApVf in the theta, alpha, and beta bands around the left occipital area.

4. Discussion

Our analyses revealed that the reported responses of the nine subjects with the same affective rates to each unimodality corresponded to the MEG activity results, with similar T-scores for each unimodality in the frontal and temporal brain areas. Moreover, their reported responses of a peaceful feeling to the bimodal AfVp and of a fearful feeling to ApVf corresponded to the MEG activity results, with a higher T-score for the difference between ApVp and ApVf and with a lower T-score for the difference between ApVp and AfVp in the frontal and temporal brain areas. These observations suggest that information extraction in a single modality can be completely dependent on this pathway [11]. However, the various types of information elicited by a multimodality from different pathways may exert competitive effects, especially in incongruent conditions, and a bias sensation that would manifest finally after fusion [3,13]. The event-related potential research of Fort et al. [18] revealed that the magnitude of early interactions in recognition partly depended on the dominant sensory modality of the subjects, with those who were considered to be auditory dominant being faster to recognize auditory objects than visual objects, and visual dominant individuals recognizing visual objects more rapidly than auditory objects. Latinus et al. [7] also demonstrated that, to congruent and incongruent bimodal stimuli, the processing of faces dominate over the processing of voices. Most of the subjects (9 out of 11) in the present study perceived more emotion to the visual stimulus in the audiovisual modality. However, our results demonstrate the presence of diversity between individuals rather than representing a universal state. The dominance can depend on the period and quality of the stimulus usage as well [9].

5. Acknowledgements

This study was supported in part by research grants from National Science Council (NSC-99-2410-H-130-018-, NSC-100-2628-E-010-002-MY3) and Academia Sinica (AS-99-TP-AC1), Taiwan.

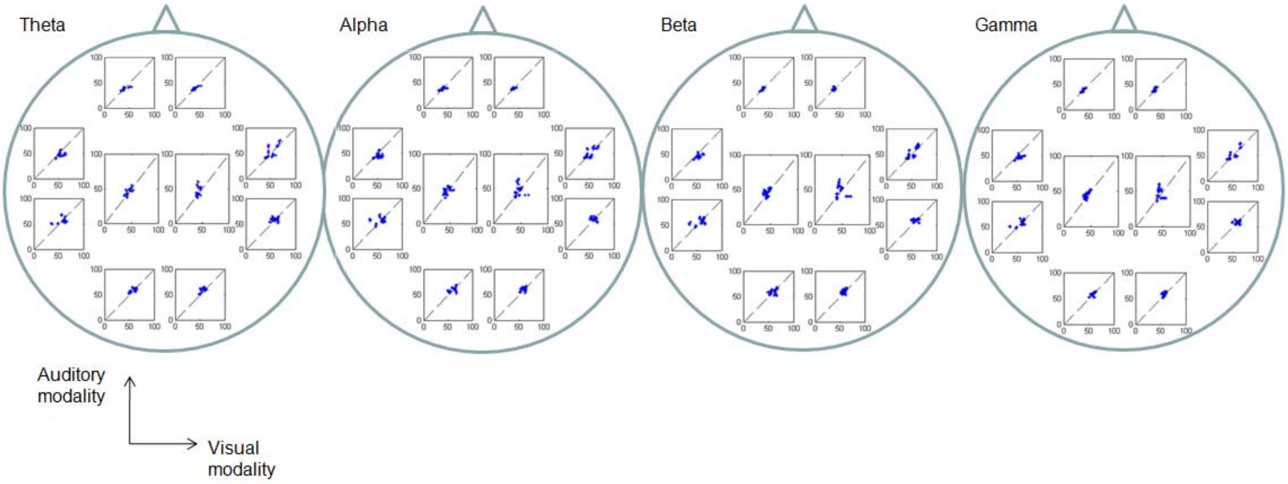

Fig. 1: Audiovisual maps of the theta, alpha, beta, and gamma activities in ten different regions from the nine subjects. Each dot represents the response at one frequency for one subject. The diagonal line indicates equal weighting (1:1) of the two modalities.

Table 1: Mean T-score ratios of unimodalities and the differences between bimodalities at four frequency bands across ten brain areas.

Modality:		mean(Ap/Vp)				mean[(ApVp – AfVp)/(ApVp – ApVf)]			
Location		Theta	Alpha	Beta	Gamma	Theta	Alpha	Beta	Gamma
Left	Frontal	0.98	0.99*	0.99*	0.98	0.96	0.96*	0.96*	0.98
	Fronto-temporal	0.88*	0.87*	0.87*	0.87*	1.00*	0.99*	1.00*	1.00*
	Temporal	1.02	1.04	1.01	1.03	1.00	0.98	1.00	1.03
	Central	1.03*	1.04*	1.03*	1.02*	0.98*	0.98*	0.98*	0.98*
	Occipital	1.02	1.02*	1.02*	1.02*	0.99	0.97*	0.97*	0.99*
Right	Frontal	0.98	0.99*	1.00*	1.00*	0.98	0.95*	0.96*	0.96*
	Fronto-temporal	1.05	1.01	1.02*	1.03*	0.99	0.96	0.96*	0.98*
	Temporal	1.01	1.01*	1.01*	1.00*	0.98	0.96*	0.96*	0.96*
	Central	1.07*	1.05*	1.05*	1.05*	0.97*	0.98*	0.98*	0.97*
	Occipital	1.05*	1.06*	1.06*	1.05*	0.98*	0.99*	0.99*	1.01*

* Significant difference ($p < 0.05$; Student's *t*-test) relative to modality.

6. References

[1] Tanaka, E., Kida, T., Inui, K., Kakigi, R., 2009. Change-driven cortical activation in multisensory environments: an MEG study. Neuroimage 48, 464–474.

[2] Logeswaran, N., Bhattacharya, J., 2009. Crossmodal transfer of emotion by music. Neurosci Lett 455, 129–133.

[3] Stekelenburg, J.J., Vroomen, J., 2009. Neural correlates of audiovisual motion capture. Exp Brain Res 198, 383–390.

[4] Doehrmann, O., Weigelt, S., Altmann, C.F., Kaiser, J., Naumer, M.J., 2010. Audiovisual functional magnetic resonance imaging adaptation reveals multisensory integration effects in object-related sensory cortices. J Neurosci 30, 3370–3379.

[5] Esposito, A., 2009. The perceptual and cognitive role of visual and auditory channels in conveying emotional information. Cognit Comput 1, 268–278.

[6] Dolan, R.J., Morris, J.S., de Gelder, B., 2001. Crossmodal binding of fear in voice and face. Proc Natl Acad Sci U S A 98, 10006–10010.

[7] Latinus, M., VanRullen, R., Taylor, M.J., 2010. Top-down and bottom-up modulation in processing bimodal face/voice stimuli. BMC Neurosci 11, 36.

[8] Schifferstein, H.N.J., Otten J.J., Thoolen F., Hekkert P., 2010. The experimental assessment of sensory dominance in a product development context. J Design Res 8, 119–144.

[9] Fenko, A., Schifferstein, H.N., Hekkert, P., 2010. Shifts in sensory dominance between various stages of user-product interactions. Appl Ergon 41, 34–40.

[10] Pourtois, G., de Gelder, B., Bol, A., Crommelinck, M., 2005. Perception of facial expressions and voices and of their combination in the human brain. Cortex 41, 49–59.

[11] Zhao, H., Chen, A.C.N., 2009. Both happy and sad melodies modulate tonic human heat pain. J Pain 10, 953–960.

[12] Pourtois, G., Debatisse, D., Despland, P.A., de Gelder, B., 2002. Facial expressions modulate the time course of long latency auditory brain potentials. Brain Res Cogn Brain Res 14, 99–105.

[13] Koelewijn, T., Bronkhorst, A., Theeuwes, J., 2010. Attention and the multiple stages of multisensory integration: A review of audiovisual studies. Acta Psychol (Amst) 134, 372–384.

[14] Brainard, D.H., 1997. The Psychophysics Toolbox. Spat Vis 10, 433–436.

[15] Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis 10, 437–442.

[16] Adachi, Y., Shimogawara, M., Higuchi, M., Haruta, Y., Ochiai, M., 2001. Reduction of non-periodic environmental magnetic noise in MEG measurement by continuously adjusted least squares method. IEEE Trans Appl Supercon 11:669–672.

[17] Hyvärinen, A., Särelä, J., Vigário, R., 1999. Spikes and bumps: artefacts generated by independent component analysis with insufficient sample size. In Proc. Int. Workshop on Independent Component Analysis and Signal Separation (ICA’99), 425–429.

[18] Fort, A., Delpuech, C., Pernier, J., Giard, M.H., 2002. Dynamics of cortico-subcortical cross-modal operations involved in audio-visual object detection in humans. Cereb Cortex 12, 1031–1039.