Audio - Visual Interaction in P300

[1]Aparna T.H. & [2]Sandeep Maruthy

Abstract

The present study investigated the interaction in the processing of auditory and visual stimuli in the generation of P300 and the effect of electrode site on P300. P300 was obtained for auditory alone condition, audio-visual congruent and audio-visual incongruent conditions, which in turn were between two electrode sites (C_z & Pz) for latency and amplitude. The experimental data was obtained from 20 normal hearing adults. The results showed that the mean latencies were shorter in audio-visual congruent condition compared to that in auditory mode and audio-visual incongruent conditions. The mean amplitude was higher in the audio-visual congruent condition compared to auditory and audio-visual incongruent conditions in both the electrode sites. The mean latency was shorter and amplitude was higher in Pz than that in Cz in all the three stimulus conditions suggesting that distribution of the electrical field and the dipole of the potential can be better recorded from the regions closer to the parietal lobe compared to mid and frontal regions of the brain. The P300 recordings supported earlier perceptual processing and better speech perception in auditory-visual mode compared to that in auditory mode suggesting that bimodal condition facilitates the detection of the stimulus.

Keywords:*P300, audio-visual interaction, congruent, incongruent.*

Introduction

Audiovisual modality for speech perception is a well known rehabilitative strategy in individuals with hearing impairment, where only auditory modality is not providing sufficient cues. the very logic of this strategy is that the bimodal stimulation facilitates speech perception compared to unimodal stimulation. Various behavioral studies have shown that behavioral responses to auditory (A) and visual (V) stimuli are improved when those stimuli are accompanied by a task-relevant stimulus in the other modality (Odgaard, Arieh & Marks, 2004; Lippert, Logothetis & Kayser, 2007).

Physiologically, visual influences in the auditory cortex would result from feedback projections from polysensory areas (Calvert, Campbell & Brammer, 2000; Miller & D'Esposito, 2005), particularly from the superior temporal sulcus. Several studies have shown that auditory event-related potentials (ERPs) can be altered by visual speech cues as early as the N1 (Besle, Fort, Delpuech & Giard, 2004; Mottonen, Schurmann & Sams, 2004; Wassenhove, Grant & Poeppel, 2005) that is, during the building of an auditory neural representation itself (Naatanen & Winkler, 1999). Such an enhancement however also depends on the attention. Talsma, Dotty and Woldorff (2007) reported that when attention was directed to both auditory and visual modalities simultaneously, audio-visual interaction occurred in early sensory processing. However when only one modality was attended, the interaction process was delayed to later in processing. Cross-modal enhancement of neural activity has been demonstrated using electrophysiological recordings (Giard & Peronnet, 1999; Foxe, et al., 2000) as well as neuro imaging measures (Calvert, et al., 2000).

Although it is clear that audio-visual integration enhances speech perception, the neuro-physiological basis of it is yet to be completely explored. The influences of multimodal stimulation on the neural processing would throw light on the time course of multisensory interaction. It would lead to understanding of whether multisensory stimulation results in speeding up the perception process, strengthen the perception or both.

The P300 being the first true event related potential, is expected to show changes with polysensory stimulation, if any. P300 is not a modality specific response and therefore can be elicited with auditory, visual as well as somatosensory stimuli. If P300 is found useful in assessing audiovisual interaction, it can be used as a valuable tool in assessing individuals with deficits in audiovisual interaction like in dyslexia, central auditory processing disorder, autism spectrum disorder and schizophrenia. Thus the present study aimed to document the characteristics of P300 in different audiovisual conditions.

The primary objective was to investigate the effects of audiovisual interaction on the latency and amplitude of P300. The secondary objective was to study the latency and amplitude differences among Cz and Pz electrode sites, if any.

Method

The study was based on the hypothesis that cross modal interaction of auditory and visual domain has an effect on the latency and amplitude of P300. The following method was adopted to test the hypothesis.

[1]Email: aputh10@gmail.com,

[2]Reader in Audiology, Email: msandeepa@gmail.com

Participants

25 years participated in the study. All the participants had puretone thresholds within 15dB HL at octave frequencies between 250 and 8000 Hz. Their speech identification scores (SIS) were normal in both quiet and in noise (SIS above 90% in quiet & above 60% in noise). All the participants had normal middle ear functioning indicated by type A tympanogram and presence of acoustic reflexes bilaterally. They had normal or corrected-to-normal vision verified using a Snell's chart at a distance of 6 feet. They did not have any past/present history of neurological problems.

The participants were screened for central auditory processing disorders using Speech in noise test (SPIN) in which they obtained a score of >60% in both ears. All the participants were meritorious students from different parts of the country pursuing their bachelor's or master's degree in Speech and Hearing.. They were blinded to the purpose of the study and a written consent was taken from each participant prior to testing. The method used in the study was approved by the institutional review board.

Stimuli and Test Environment

Auditory as well as visual stimuli were used for recording P300. The stimuli consisted of two CVC words; /cat/ and /bat/. These two CVC words were chosen as they were minimal pairs and were picturable. A minimal pair was preferred as the two words of the pairs would elicit similar LLRs, which in turn was necessary for better P300 detectability. The pictures of the cat and bat were used as visual stimuli. The words were spoken by an adult native Kannada speaker and were recorded by a unidirectional microphone in a sound treated room. Stimuli were digitally recorded using Praat Software (version 5.1.31) at a sampling frequency of 44,100 Hz and 16 bit digitization. The stimuli were then edited for noise and hiss reduction using Adobe Audition (Version 1.5). They were normalized to the same scaling factor and were edited to restrict the duration to 250 ms.

The two auditory stimuli were presented in Odd ball stimulus paradigm. Stimulus /cat/ was presented frequently while the stimulus /bat/ was presented infrequently. In the audio-visual mode of presentation, visual stimuli which included pictures of /cat/ and /bat/ were presented only with infrequent auditory stimulus-/bat/. Depending on the visual stimuli presented, audio-visual mode was either termed as congruent or incongruent. If picture of /bat/ was presented along with the auditory infrequent /bat/, it was AV- congruent condition. On the other hand, if picture of /cat/ was presented along with the auditory infrequent /bat/, it was AV-incongruent condition. The oddball paradigm involved the presentation of one infrequent after two frequent stimuli.

The auditory stimulus was 250 ms in duration while the visual stimulus was 1000 ms in duration. The visual stimulus was triggered along with the auditory stimuli through an external laptop during audio-visual presentation. A laptop computer was used for visual stimulus presentation. Recording of the stimulus and all the evaluations were performed in sound treated rooms where the noise levels were within permissible limits (ANSI.S3.1, 1991). The rooms were also electrically insulated.

Recording of P300

To record P300, the participants were made to sit in a comfortable reclining chair and were asked to relax. The Cz and Pz electrode sites were cleaned with skin preparation gel and the disc electrodes were placed using a conduction paste. Prior to recording P300, an absolute impedance of less than 5 kOhms and relative impedance of less than 2 kOhms was ensured. The auditory stimuli were presented through insertphones at 80 dBnHL. The repetition rate used for auditory stimulus was 1/s and 1/3s for visual stimulus. The resultant auditory evoked potentials were recorded from two scalp electrode sites (Cz & Pz) referenced to the nose tip. The recorded responses were then amplified (50,000 times) and band pass filtered between 1 and 30 Hz. The responses were averaged totally for 300 stimuli and were analyzed for 750 ms.

The participants were instructed to minimize eye blinks to reduce the contamination of the desired EEG. They were instructed to pay attention to the blocks of stimuli which were presented and were asked to mentally count the infrequent stimulus (bat) during the auditory presentation mode. During audio-visual presentation mode, the participants were asked to press an arrow key in a computer keyboard, after hearing each stimulus and to watch the monitor kept in front of participant which displayed the pictures. The visual stimulus would arrive along with infrequent auditory stimulus, and was triggered by the arrow key pressed for the second frequent stimulus. The visual stimulus, although triggered by the second frequent, was displayed along with infrequent stimulus only. P300 was recorded for 3 different stimulus paradigms: The 3 stimulus paradigm included, auditory, audio-visual congruent and audio-visual incongruent. The order of the 3 paradigms was randomly used.

Analyses

The P300 was identified in each participant, for each stimulus paradigm, and at each electrode site. The response was analyzed to note down onset latency, peak latency, offset latency and the peak amplitude. The responses were subjectively analyzed by two experienced audiologists and the average waves recorded for the frequent and infrequent stimuli were compared. P300 was detected in the infrequent wave. The criteria for on-

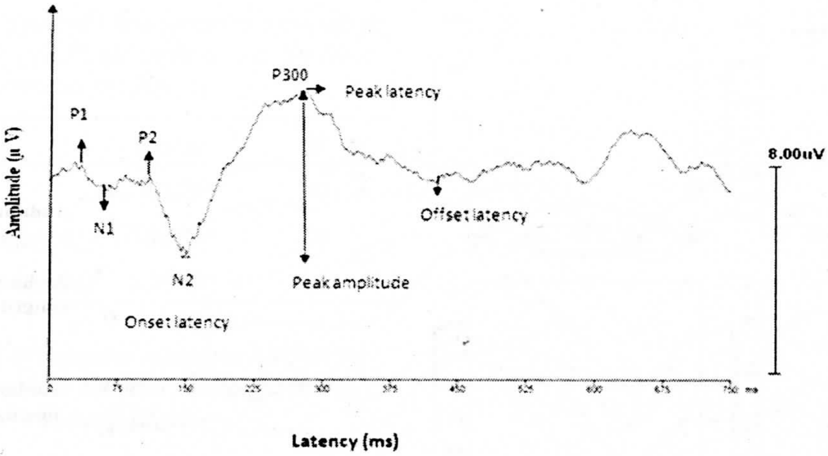

Figure 1: A representative wave recorded in auditory modality with the target parameters marked.

set and offset latency were set based on grand averaged waveforms of infrequent stimulus in each modality. A representative recording with target parameters marked is shown in Figure 1.

Results

The primary aim of the present study was to analyze the effects of audio-visual interaction on P300 latency and amplitude. The secondary aim was to evaluate the effect of electrode site on P300 latency and amplitude. While statistically testing these effects, presentation mode and electrode site were treated as independent variables, while amplitude and latency of P300 served as dependent variables. The statistical analysis was carried out using Statistical Package for Social Sciences (SPSS version 16).

Effect of Electrode site on P300

The effect of electrode site was tested for both latency and amplitudes of P300. The results are reported separately for the 2 parameters. There were 3 latency parameters measured in the P300 waves: onset latency, peak latency and offset latency. The effect of electrode site was analyzed for all the 3 parameters. Figure 2 shows the mean and standard deviation of the onset latency, peak latency and the offset latency recorded at the Cz and Pz electrode sites.

The data showed that the mean latencies at Cz were prolonged compared to that at Pz. This was true in all the three modalities and for all the three latency parameters. The observed mean differences were then tested for statistical significance on paired t-test. The results of paired t-test (Table 1) showed that there was a significant difference between the latencies recorded from the 2 electrode sites. This was true in all the three stimulus conditions. Figure 3 represents the waveforms for three stimulus conditions across the two electrode sites.

The mean and standard deviation of the peak amplitude of P300 are given in Figure 4. The mean amplitudes in Pz were higher than that in Cz in all the three stimulus conditions. The comparison of mean amplitudes between the two electrode sites was done on paired t-test separately for each stimulus conditions. Results (Table 2) showed significant difference between the two electrode sites in their mean amplitudes, in all the three stimulus conditions.

Effect of Stimulus Condition on Latency and Amplitude of P300

The mean and standard deviation of onset latency, peak latency and offset latency in the three stimulus conditions are shown in Figure 2. The effect of stimulus condition on the latency parameters was analyzed separately for the data of Cz and Pz electrode sites. In general the mean latencies were shorter in audio-visual congruent condition compared to that in auditory

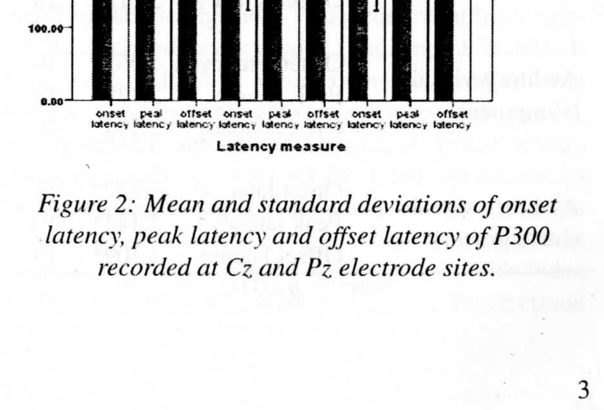

Figure 2: Mean and standard deviations of onset latency, peak latency and offset latency of P300 recorded at Cz and Pz electrode sites.

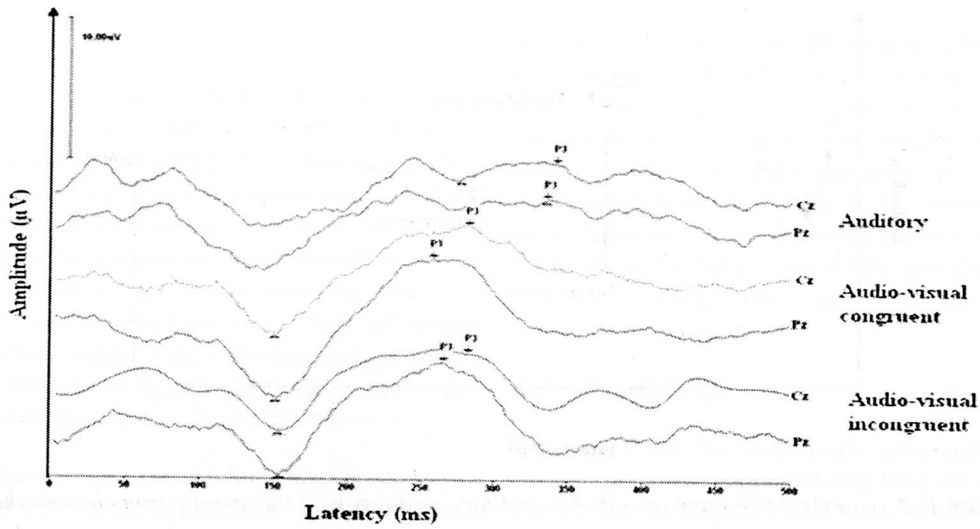

Figure 3: Waveforms for three stimulus conditions at two electrode sites.

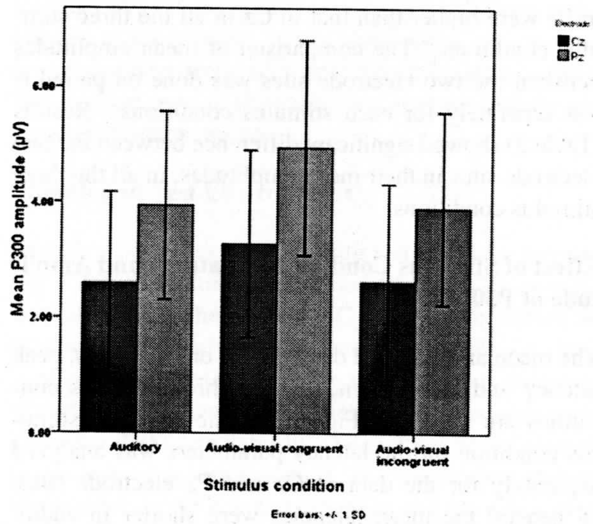

Figure 4: Mean and standard deviation of the peak amplitude of P300 in the 3 stimulus conditions.

Table 1: Results of paired t-test comparing the onset latency, peak latency and offset latency recorded from Cz and Pz electrode sites in the three stimulus conditions

Conditions	Parameters	t value	df
Auditory	Onset latency	6.222*	19
	Peak latency	1.823*	19
	Offset latency	2.618*	19
Auditory-visual (Congruent)	Onset latency	4.062*	19
	Peak latency	2.779*	19
	Offset latency	3.544*	19
Auditory-visual (Incongruent)	Onset latency	3.237*	19
	Peak latency	4.444*	19
	Offset latency	4.309*	19

*Note: *- p<0.05*

and audio-visual incongruent conditions. The effect of stimulus condition was tested on Repeated measures ANOVA and the results (Table 3) of Cz site showed a significant main effect of stimulus condition on only onset latency. The effect was absent on peak latency and offset latency. On the other hand at the Pz electrode site, there was significant main effect of condition on all the three latency parameters.

Wherever there was a significant main effect of condition, pair-wise comparison was tested using Bonferroni post hoc test. The results of the Bonferroni post hoc test showed that, for the data of Cz electrode site there was a significant difference between the onset latency recorded in the auditory and the audio-visual congruent condition. However there was no significant difference in the onset latency between auditory and audio-visual incongruent and, audio-visual congruent and audio-visual incongruent condition.

Similar pattern of results as in Cz electrode site was obtained at Pz site with a significant difference between the onset latency, peak latency and offset latency recorded in the auditory and the audio-visual congruent condition. However there was no significant difference in the onset latency and peak latency between auditory and audio-visual incongruent and audio-visual congruent and audio-visual incongruent conditions.

The mean and standard deviation of the amplitude recorded in the three stimulus conditions at the two electrode sites are shown in Figure 3. The mean amplitude was higher in the audio-visual congruent condition compared to auditory and audio-visual incongruent conditions. The mean data was similar in the auditory and audio-visual incongruent condition. The same trend was observed at both the electrode sites. Repeated measures ANOVA was used to test the effect of stimulus condition on amplitude and the results showed a signif-

Table 2: Results of paired t-test comparing amplitude recorded from Cz and Pz electrode sites in the three stimulus conditions

Conditions	t value	df
Auditory	-10.75*	19
Audio-visual congruent	-6.523*	19
Audio-visual incongruent	-7.556*	19

Table 3: Results of Repeated measures ANOVA for latencies across three stimulus conditions at two electrode sites

Electrode site	Parameters	F	df (error)
Cz	Onset latency	11.32*	2 (18)
	Peak latency	2.456	2 (18)
	Offset latency	1.227	2 (18)
Pz	Onset latency	11.872*	2 (18)
	Peak latency	3.487*	2 (18)
	Offset latency	3.716*	2 (18)

*Note: *- p<0.05*

Table 4: Correlation coefficient showing the relationship between index I and II derived for onset latency, peak latency, offset latency and peak amplitude

Electrode Site	Parameters	r
Cz	Onset latency	0.474*
	Peak latency	0.301
	Offset latency	0.517*
	Amplitude	0.809**
Pz	Onset latency	0.261
	Peak latency	0.481*
	Offset latency	0.444
	Amplitude	0.880**

*Note: *- p<0.05, **- p¡0.01*

icant main effect of stimulus condition on amplitude of P300 at Cz [F(2,18) = 4.291, p = 0.021] and Pz [F (2,18) = 7.208, p = 0.002].

As there was a main effect of stimulus condition, the pair-wise comparison was tested using Bonferroni post-hoc test. The results showed that audio-visual congruent condition was significantly different from auditory as well as in audio-visual incongruent condition. However there was no significant difference in amplitude between auditory and audio-visual incongruent condition. The results were same for both Cz and Pz.

Correlation of Latency Index and Amplitude Index

Results in the previous sections showed that the effect of stimulus condition was present on both latency and amplitude of P300. However, it did not give the picture of the effects of two audio-visual conditions (congruent & incongruent) related with each other.

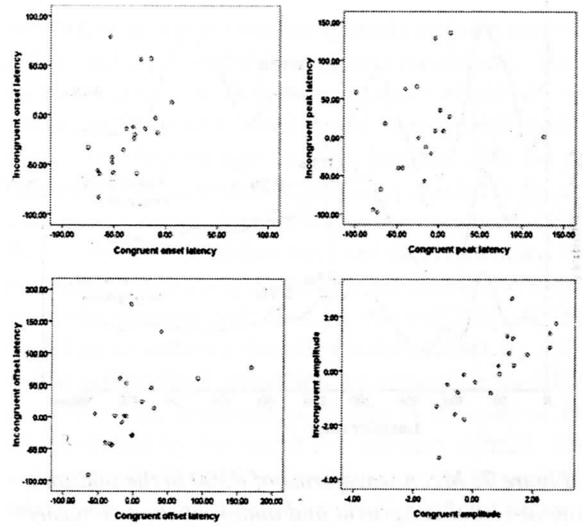

Figure 5: Scatter plot showing the relationship between Index I and Index II derived for onset latency, peak latency, offset latency and peak amplitude at Cz site.

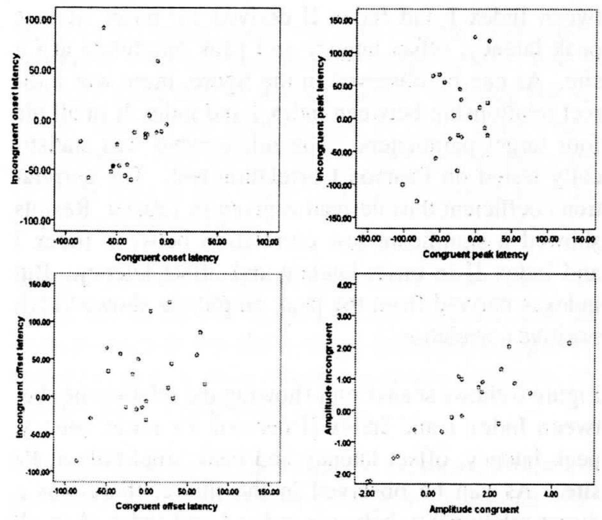

Figure 6: Scatter plot showing the relationship between Index I and Index II derived for onset latency, peak latency, offset latency and peak amplitude at Pz site.

To derive this relationship, latency and amplitude indexes were determined. This was done by taking the difference in the latency obtained in audio-visual conditions and auditory mode. The difference values obtained by subtracting audio-visual congruent from auditory were termed, latency index I and amplitude index I. Whereas the difference values obtained by subtracting audio-visual incongruent from auditory were termed, latency index II and amplitude index II. These indexes were calculated separately for the data from Cz and Pz electrode sites. To see the relation between the changes seen due to audio-visual congruent and AV-incongruent stimulus paradigms, Index I was correlated with Index II. This was done separately for Cz and Pz electrode site.

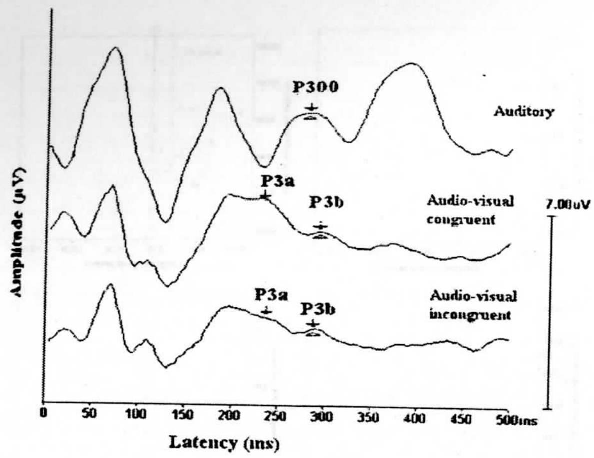

Figure 7: Mean waveforms of P300 in the auditory, audio-visual congruent and audio-visual incongruent conditions showing difference in the morphology. Auditory P300 shows single peak while audio-visual condition shows bifid peak.

Figure 5 shows scatter plot showing the relationship between Index I and Index II derived for onset latency, peak latency, offset latency and peak amplitude at Cz site. As can be observed in the figure, there was a direct relationship between index I and index II in all the four target parameters. The relationship was statistically tested on Pearson Correlation test. The correlation coefficient thus derived is given in Table 4. Results showed a significant low correlation between index I and index II in onset latency and offset latency. But indexes derived from the peak amplitude showed high positive correlation.

Figure 6 shows scatter plot showing the relationship between Index I and Index II derived for onset latency, peak latency, offset latency and peak amplitude at Pz site. As can be observed in the figure, there was a direct relationship between index I and index II in all the four target parameters. The relationship was statistically tested on Pearson Correlation test, and the results are given in Table 4. Results showed a significant low correlation between index I and index II for peak latency. But indexes derived from the peak amplitude showed high positive correlation. There was no significant correlation between index I and index II for onset latency and offset latency.

Morphological Changes in P300

For all the subjects, better waveforms were obtained at Pz compared to Cz. Comparison of P300 waveforms across auditory condition, audio-visual congruent and audio-visual incongruent conditions showed that the amplitude of P300 was larger in audio-visual conditions compared to auditory condition. It was observed that in auditory condition, a single P300 peak was seen. Whereas bifid peaks (P3a & P3b) were present in audio-visual conditions. Broadening of the P300 peak was noticed in audio-visual condition when compared to auditory condition. Figure 7 represents the mean waveforms of P300 depicting morphological changes in P300 across the three stimulus conditions.

Discussion

P300 is an endogenous potential recorded for an odd ball paradigm. Considering that the potential is generated by auditory association areas, it was hypothesized in the present study that P300 shall differ in bimodal stimulation compared to unimodal stimulation. Further, the effect of electrode site on P300 was analyzed. The analysis of independent effects of the 2 variables; mode of stimulation and the electrode site on the latency and amplitude of P300 showed some interesting findings. Attempt has been made to justify these findings and derive appropriate implications in the following sections.

Effect of Electrode Site on P300

The present study revealed that longer latencies (onset, offset & peak latency) and reduced amplitude at Cz than Pz. It is evident from this result that spatial amplitude distribution of P300 is largest at the parietal electrode sites and is attenuated as the recording site move to central locations. The exact source of the potential, region of integration and the principal components contributing cannot be inferred due to the reduced number of electrodes used in this study. However, the findings suggests that the distribution of the electrical field can be better recorded from the regions closer to the parietal lobe compared to mid and frontal regions of the brain.

These observations are in accordance with various earlier studies. Katayama and Polich (1998) obtained P300 using three stimulus odd ball paradigms and found that P300 target amplitude was larger across the midline electrode sites, being largest over the parietal site. Polich (2007) assessed correlational association between peak amplitude and latency for the P300 using auditory discrimination task. They found that for the target stimuli, P300 amplitude and latency were negatively correlated over the frontal-central and right medial/lateral recording sites. Similar results were reported by Comerchero and Polich (1999). In these studies larger P300 amplitude have been recorded at parietal sites.

Effect of stimulus conditions on latency and amplitude of P300

As evident from the results of the present study at Pz, all the 3 latency parameters (onset latency, peak latency & offset latency) were shorter in audio-visual congruent condition compared to auditory mode and audio-visual incongruent conditions. This is an electrophysiological evidence of the earlier onset of neural processing pertaining to sound discrimination, underlying

P300. Functionally, this may mean earlier perceptual processing in audio-visual congruent condition suggesting that bimodal condition accelerates the detection of the stimulus. However, this needs to be experimentally confirmed through correlation of behavioral and electrophysiological findings.

These findings are consistent with the findings of Fort, Delpuech, Pernier and Giard (2002). They assessed the effect of unimodal (auditory or visual) and non redundant bimodal conditions (auditory & visual) on ERPs. Event-related potential analysis revealed the existence of early (200 ms latency) cross modal activities in sensory specific and nonspecific cortical areas suggesting cross modal interaction resulting in facilitation effect (shorter reaction time) for the detection of bimodal stimuli compared to unimodal stimuli. Similar findings were reported in various earlier studies (Odgaard et al., 2004; Frassinetti, Bolognini & Ladavas, 2002).

Behavioral studies (Sumby & Pollack, 1954; O'Neil, 1954; Erber, 1969; Grant & Seitz, 2000., Rudmann, McCarley & Kramer, 2003; Bernstein, Auer & Takayanagi, 2004; Ross, Saint-Amour, Leavitt, Javitt & Foxe, 2007) unanimously show better speech perception in auditory-visual mode compared to that in auditory mode. The earlier latency of P300 may be the neurophysiological basis for the behavioral advantage in AV mode. Auditory-cognitive processing of information is expected to be better in auditory-visual congruent condition since the system is able to utilize the information from both the modalities in a complementary manner.

There was no significant difference in latencies (both in Cz and Pz) between auditory and audio-visual incongruent condition, also between audio-visual congruent and audio-visual incongruent condition. The exact underlying physiological reason could not be derived for this result. Logically, it was expected that the AV incongruent condition would be delayed than that in auditory mode, as the incongruency between the 2 modalities would create confusion. But the absence of the difference between the latencies of 2 stimulus conditions indicates that the processing is delayed by the incongruency. This probably is because, in instances of incongruency, the processing may be primarily determined by the dominant modality, which is auditory in this case.

The P300 picked up at Cz however showed the same trend as at Pz but only in onset latency. This could be due to differential contributions of the latent components of P300 at the 2 sites. That is, some of the latent components of P300 picked up at Pz are not picked up at Cz. Unlike at Pz, the latent components contributing for peak latency of the response at Cz were probably not sensitive to the differences in the mode of stimulation.

The present study indicated that the mean amplitude was higher in audio-visual congruent condition than auditory and audio-visual incongruent conditions. There was no significant difference in mean amplitude between auditory and audio-visual incongruent conditions. The finding again can be justified with the integration of the complimentary cues provided in the auditory and visual modes. Because the cues provided in the incongruent condition were not complimentary, enhancements in the amplitude were not seen in the audio-visual incongruent condition. Li, Wu and Touge (2010) investigated auditory detection enhancement by cross modal audio-visual interaction, by comparing the ERPs elicited by the audio-visual stimuli to the sum of the ERPs elicited by the visual and auditory stimuli. Results suggested that behavioral detection of an auditory stimulus is enhanced by the interaction of auditory and visual stimuli around 300ms in polysensory regions of the brain.

Relation between Changes Seen in Congruent and Incongruent Conditions

Because the mean changes in audio-visual incongruent condition were not significantly different from that in auditory condition, it was of interest to study the relation between the changes seen in the 2 audio-visual conditions. This would help in inferring whether the changes seen in the audio-visual incongruent were facilitative or deleterious.

Results showed a significant positive correlation between the 2 indices derived. This supports that the changes seen in the audio-visual incongruent were facilitative, like in audio-visual congruent condition, although not to a noticeable extent.

Overall, from the findings it can be derived that bimodal presentation, neurophysiologically, is beneficial over unimodal presentation for speech processing. Within bimodal conditions, congruent condition is more facilitative compared to incongruent condition and P300 can evidence the modality-based changes in the cortical auditory processing.

Conclusions

From the findings of the study it can be inferred that the distribution of the electrical field and the dipole of the potential can be better recorded from the regions closer to the parietal lobe compared to mid and frontal regions of the brain. The P300 recordings supported earlier perceptual processing and better speech processing in auditory-visual mode compared to that in auditory mode suggesting that bimodal condition accelerates the detection of the stimulus.

The P300 picked up at Cz however showed the same trend as in Pz but only in onset latency indicating that the latent components contributing for peak latency of the response at Cz were probably not sensitive to the differences in the mode of stimulation. But, in instances of incongruency, the processing may be primarily determined by the dominant modality, which is auditory in this study. Finally, it is concluded that both changes seen in audio-visual congruent and audio-visual incongruent conditions are facilitative but in audio-visual congruent condition the facilitation in evident.

The findings of the present study have important implications in diagnostic as well as rehabilitative audiology. The present study proved that P300 can be an objective index of facilitation by bimodal presentation (audio-visual). Hence it can be used in the assessment of cross modality integration at the cortical level in individuals with auditory processing disorders. P300 can also be used to monitor the progress secondary to training of cross modality integration.

References

American National Standard Institute (1991). *Specification for Maximum permissible ambient noise levels (MPANLs) allowed in an audiometric test room.* ANSI S3.1-1999 (R 2003). New York: American National Standard Institute.

Bernstein, L. E., Auer, E. T., & Takayanagi, S. (2004). Auditory speech detection in noise enhanced by lipreading. *Speech Communication, 44*(1), 5-18.

Besle, J., Fort, A., Delpuech, C., & Giard, M. H. (2004). Bimodal speech: Early suppressive visual effects in the human auditory cortex. *European Journal of Neuroscience, 20,* 2225-2234.

Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in human heteromodal cortex. *Current Biology, 10,* 649-657.

Comerchero, M.D., & Polich, J. (1999). P3a and P3b from typical auditory and visual stimuli. *Clinical Neurophysiology, 110,* 24-30.

Erber, N. (1969). Interaction of audition and vision in the recognition of oral speech stimuli. *Journal of Speech and Hearing Research, 12,* 423-425.

Frassinetti F., Bolognini N., & Ladavas E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Experimental Brain Research, 147*(3), 332-43.

Fort, A., Delpuech, C., Pernier, J., & Giard M. (2002). Dynamics of Cortico-subcortical Cross-modal Operations Involved in Audio-visual Object Detection in Humans. *Cerebral Cortex, 12*(10), 1031-1039.

Foxe, J. J., Morocz, I. A., Murray, M. M., Higgins, B. A., Javitt, D. C., & Schroeder, C. E. (2000). Multisensory auditory-somatosensory interactions in early cortical processing revealed by high-density electrical mapping. *Cognitive Brain Research, 10,* 77-83.

Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *Journal of Cognitive Neuroscience, 11,* 473-490.

Grant, K. W., & Seitz, P. F. (2000). The recognition of isolated words and words in sentences: Individual variability in the use of sentence context. *Journal of the Acoustical Society of America, 107,* 1000-1011.

Katayama, J., & Polich, J. (1998). Stimulus context determines P3a and P3b. *Psychophysiology, 35,* 23-33.

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology, 38,* 557-577.

Li, Q., Wu, J., & Touge, T. (2010). Audio-visual interaction enhances auditory detection in late stage: an event-related potential study. *NeuroReport, 21,* 173-178.

Lippert, M., Logothetis, N. K., & Kayser, C. (2007). Improvement of visual contrast detection by a simultaneous sound. *Brain Research, 1173,* 102-109.

Miller, L. M., & D'esposito, M. (2005). Perceptual Fusion and Stimulus Coincidence in the Cross-Modal Integration of Speech. *The Journal of Neuroscience, 25,* 5884-5893.

Mottonen, R., Schurmann, M., & Sams, M. (2004). Time course of multisensory interactions during audio-visual speech perception in humans: a magnetoencephalographic study. *Neuroscience Letters, 363,* 112-115.

Naatanen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin, 125,* 826-859.

Odgaard, E. C., Arieh, Y., & Marks, L. E. (2004). Brighter noise: sensory enhancement of perceived loudness by concurrent visual stimulation. *Behavioral Neuroscience, 4,* 127-132.

O'Neill, J. J. (1954). Contributions of the visual components of oral symbols to speech comprehension. *Journal of Speech & Hearing Disorders, 19,* 429-439.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology, 118*(10) 2128-2148.

Ross, L. A., Saint-Amour, D., Leavitt, V., Javitt, D. C., & Foxe J. J. (2007). Do you see what I'm saying? Optimal Visual Enhancement of Speech Comprehension in noisy environments. *Cerebral Cortex, 17*(5), 1147-53.

Rudmann, D. S., McCarley, J. S., & Kramer, A. F. (2003). Bimodal display augmentation for improved speech comprehension. *Human Factors, 45,* 329-336.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise, *The Journal of the Acoustical Society of America, 26,* 212-215.

Talsma, D., Doty, T. J., & Woldorff, M. G.(2007). Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? *Cerebral Cortex, 17,* 691-701.

Wassenhove, V., Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences of the United States of America, 102,* 1181-1186.