Introduction:

In recent years, the Sound-Induced Flash Illusion (SIFI) has become a valuable tool for investigating audiovisual (AV) processing in the brain (Hirst et al., 2020; Keil, 2020). SIFI demonstrates that sound can influence visual perception, similar to how visual lipreading alters auditory perception in AV speech (as seen in the McGurk effect) (McGurk ref). Previously, the popular view among scholars was that information-processing circuits for different senses operate independently within the brain up to the cortex and vision dominates other senses (two ref). However, SIFI and other instances of multisensory interaction suggest that auditory might dominates in some instances as in SIFI. When different types of stimulation occur at the same time, there may be neural components in the brain associated with multi-sensory interaction.

This study aims to develop electrophysiological analysis methods to find specific neural interaction components (NICs) between evoked potentials (EPs) generated by cross-modal visual (F, flash) and auditory (B, beep) stimuli.

Figure 1. The experiment set-up and the two-channel electrodes montages to collect auditory and visual EEG signals. The auditory channel is composed of Cz as the positive electrode and right and left mastoids (linked) as the negative electrode. It collects signals from the ear (which receives tone stimuli) to the auditory cortex (which processes auditory perception). The visual channel consists of Oz as the positive electrode and Fz as the negative electrode. It collects signals from the eyes (which receive flash stimuli) to the visual cortex in the occipital lobe (which processes visual perception). Fpz serves as the ground electrode to reduce the interference of external noise. In the SmartEP software (IHS, 2024), the data is recorded in real-time and the EPs are averaged every 500ms(?) to obtain the average pattern of EEG signals under different conditions.

Figure 2. Audiovisual Summation Equations (A) and Basic Audiovisual Interaction (AVI) Equations (B). The SIFI experimental visual and auditory stimuli were divided into sub-components, and the theoretical computed EPs are obtained by adding the subcomponent signals. By subtracting computed values from the real data, we may find evidence of AV interaction.

Methods:

We recorded auditory evoked potentials (AEPs) from electrodes placed at Cz (+) and linked mastoids (-), and visual evoked potentials (VEPs) from electrodes at Oz (+) and Fz (-) (Fig. 1). The Fpz electrode was grounded to minimize environmental noise (Fig. 1). Thus, all AEPs are measured from medial parietal lobe, while all VEPs are measured from occipital lobe. The SIFI program was configured in SmartEP data collection software to present subjects with nine random combinations of flash and beep stimuli (0, 1, or 2 times for one short routine), asking them to identify the stimuli they perceive. The experiment has 7 subjects. EEG data were collected during each trial, with a total of 256 sweeps and a stimulus presentation frequency of 2 Hz(?). SmartEP averages the evoked potentials every 600 ms (?) to reduce noise. After acquiring the raw data, we analyzed it using audiovisual summation equations (Fig. 2A) and basic audiovisual interaction (AVI) equations (Fig. 2B). We performed microvolt-weighted calculations of EPs using SmartEP's built-in addition and subtraction functions and marked the latency and amplitude of peaks using the labeling function.

Figure 3 – The population-averaged EPs of 7 subjects under the condition of one flash (F1), two flashes (F2), one beep (B1), two beeps (B2), and Control (no beeps and no flashes). These EPs can then be used for subsequent calculations of computed summation values. The left part of the screen is the visual EPs, and the right part is the auditory EPs.

Figure 4 – Population-averaged EPs of 7 subjects under the condition of one flash one beep (F1B1), two flashes 1 beep (F2B1), one flash two beeps (F1B2), and also two flashes two beeps (F2B2). The left part of the screen is the visual EPs, and the right part is the auditory EPs. They can then help calculate the Audiovisual Interaction.

Result:

We population-averaged the datasets from 7 subjects to obtain average EP patterns for visual and auditory channels under four types of single stimuli, four mixed audiovisual stimuli, and control conditions (Fig. 3, 4). Additionally, we analyzed the EEG data of each individual subject based on the previously mentioned equations. Figure 5 illustrates the calculation, comparison, and analysis of EPs from one subject under the F1B1 (Fig. 5A) and F2B1 (Fig. 5B) conditions.

In the average EP diagrams, we observed that the visual and auditory channels exhibit distinct neural responses under different experimental conditions. Several classical potentials associated with visual and auditory perception, such as N75, P100, and N135, were consistently observed. Furthermore, specific potentials also appeared consistently in individual data analyses. The EP analysis revealed that the net difference between the actual EP and the calculated EP is not a straight line (Fig. 5), indicating that audiovisual interaction (AVI) is unlikely to be zero. In comparing the differences between the two experimental groups and the control group, several early peaks were clearly identified. To further investigate how these data differences relate to AVI, we need to label the data and examine latency and amplitude.

Figure 5. Examples of computation and analysis of EPs of one subject Under F1B1 (A) and F2B1 (B) conditions. From top to bottom are the EPs of two subcomponents, the real EPs from experiment, computed EPs, the net difference between the two, and the difference between the two (Brown) compared with the control EPs (Orange).

Conclusion:

Using electrophysiological analysis methods, we can preliminarily confirm the existence of components that are independent of the corresponding VEP and AEP summation during the perception of combined audiovisual stimuli. These may represent the AVI components we are investigating. However, due to the small sample size, we must collect EEG data from additional subjects to confirm the generalizability of these components.

To further explore the characteristics of these peaks, we need to label the EPs and examine their latency and amplitude. Nonetheless, the current findings enhance our understanding of audiovisual sensory processing. Further research on AVI could lead to the development of new treatment options for patients with visual or auditory disorders.