Temporal integration of faces and vocalizations in the ventrolateral prefrontal cortex of non-human primates

Jaewon Hwang[1], and Lizabeth M. Romanski[2]

[1]Brain & Cognitive Science, Univ. of Rochester, Rochester, NY; [2]Dept. Neurobiology & Anatomy, Univ. of Rochester, Rochester, NY

692.04

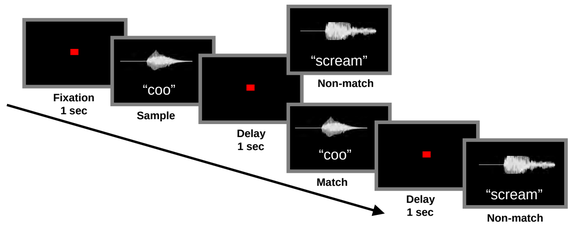
Introduction

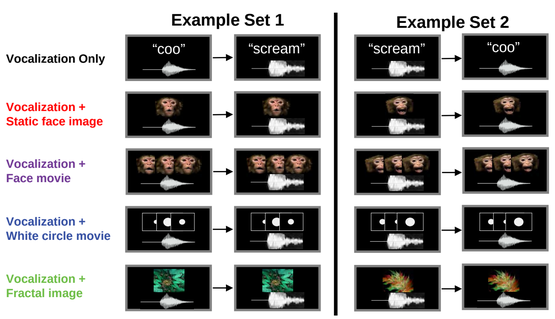
- During human speech and non-human primate vocalizations, a facial gesture (facial posture shift or mouth opening) generally precedes the generation of a vocalization.
- In addition, the propagation delay of the face and vocalization information in the brain is different. These are the temporal factors which affect face-vocalization integration.
- To examine the temporal integration between faces and vocalizations, we trained non-human primates to perform a vocalization discrimination task with and without accompanying visual stimuli and recorded single-neuron activity from VLPFC, where multisensory neurons have been previously recorded.

Methods

Non-match to Sample Vocalization Discrimination Task

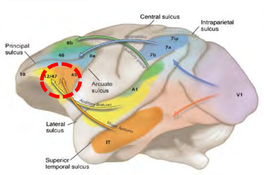
- Species-specific vocalizations and corresponding face stimuli from our home colony were used as stimuli.
- The subject was required to press a button when the vocalization non-match occurred. Correct presses were rewarded with a drop of juice.
- Occurrence of the non-match was unpredictable so that in some trials the sample was repeated and then a non-match stimulus occurred while in other trials the non-match occurred immediately after the sample.
- In our task, 5 task conditions occurred where a vocalization was presented alone or one of four different types of visual stimuli accompanied the vocalization. These visual stimuli were irrelevant to the task.





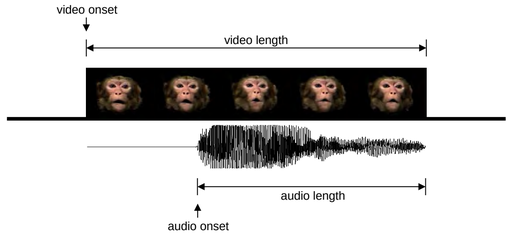
Single-neuron recording

- The recording cylinder was placed over the lateral prefrontal cortex (areas 12, 45 and 46).
- Single-neuron activity was discriminated and isolated using a signal processing system (Pentusa system, Tucker-Davis Technology, FL).



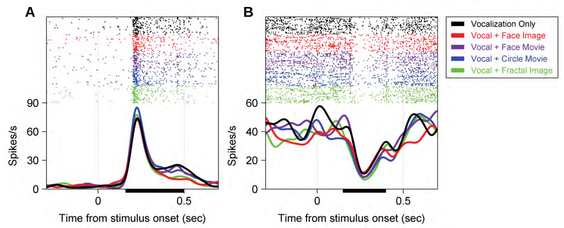
Results

Time course of the vocalization and the visual stimuli



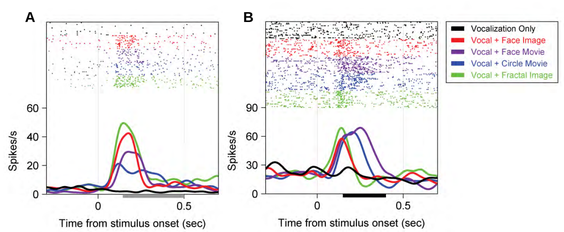
- Our face movies were shortened from the mouth opening to the vocalization offset so that only the relevant face motion and vocalization were included.
- A synthetic motion stimulus, an expanding/contracting white circle movie imitated the mouth motion and retained the same timing as the face movies.
- All the visual stimuli were presented according to the onset and the duration of the natural face-vocalization movies.

Type 1: Neurons responding to the vocalizations



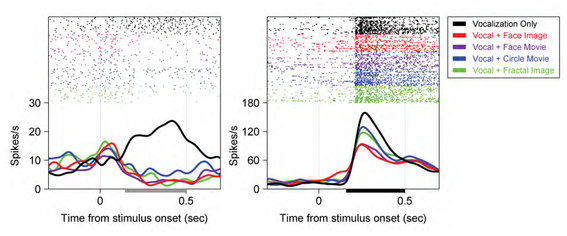
- These two example neurons responded to all stimulus conditions in the same fashion, which indicates that they responded to the common component of the stimuli, the vocalization.

Type 2: Neurons responding to the visual stimuli



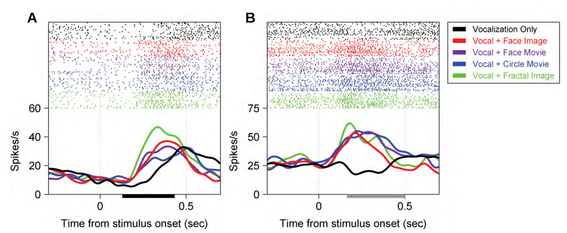
- These two neurons showed no response to the vocalization, but showed a differential activity to the conditions in which visual stimuli were presented with the vocalization, suggesting that they are unimodal visual neurons.

Type 3: Neurons showing multisensory suppression



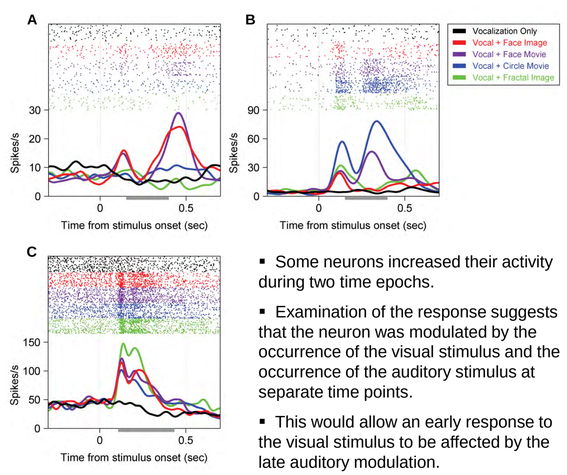
- For both of these the combination of the visual stimuli with the vocalization resulted in complete or partial suppression of the neural response.

Type 4: Neurons showing multisensory enhancement



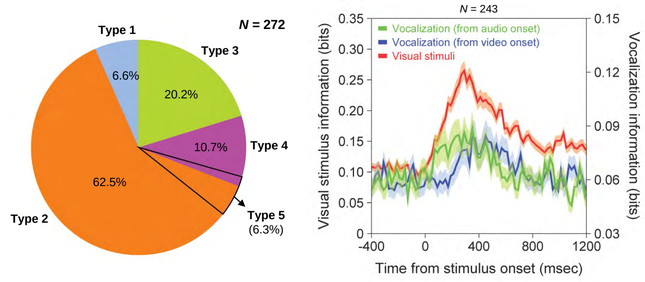
- In these neurons, presentation of the visual stimuli enhanced the response to the vocalization.

Type 5: Neurons with a biphasic response



- Some neurons increased their activity during two time epochs.
- Examination of the response suggests that the neuron was modulated by the occurrence of the visual stimulus and the occurrence of the auditory stimulus at separate time points.
- This would allow an early response to the visual stimulus to be affected by the late auditory modulation.

Population Summary



- The left panel shows the proportion of each neuron type across the population. The proportions were computed from a latency analysis and a regression analysis. Two sample stimuli were treated as separate cases.

Neural Response = a0 + a1 * **Face Image** + a2 * **Face Movie** + a3 * **Circle Movie** + a4 * **Fractal Image**

- The right panel shows the time courses of the signals related to the vocalizations and the visual stimuli in which mutual information was estimated.
- The vocalization information reached its plateau before the peak of visual information when it was computed from the audio onset, which indicates that the propagation delay of vocalization information may be shorter than that of the visual information overall.
- When vocalization information was computed with the natural timing of the stimuli, however, the latency of the signal related to the vocalizations seems to be longer than that of the visual stimuli.

Conclusions

1. The stimulus timing and the neuronal processing time of the vocalizations and the visual stimuli are two important temporal factors which affect the occurrence of audiovisual integration in VLPFC. Depending on the relative difference in these factors, neurons in the VLPFC may show different types of multisensory interactions as has been suggested for multisensory neurons in other brain regions.
2. The biphasic response observed in some VLPFC neurons suggests that the processing of vocalizations and of visual stimuli may occur sequentially in time. This may explain why the mismatched face movie changes the auditory perception in the demonstration of the McGurk effect.
3. Although the response latency for auditory stimuli is generally shorter than visual response latency, when the stimuli are presented in their natural time sequence, the vocalization information seems to affect VPLFC neurons later than the visual stimulus information.

Acknowledgments

NIDCD 04845, Cure Autism Now, and NIDCD P30 award for the Center for Navigation and Communication Sciences