Comparison of face and non-face stimuli in an audio-visual discrimination task

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

578.5

Introduction

The integration of face and vocal information is crucial to social communication. Previous studies have shown that dynamic faces elicit greater responses and may provide more information during communication than static faces. However, how the brain combines dynamic facial motion with complex vocalization stimuli is not well understood yet. To investigate these issues, we trained non-human primates to perform a vocalization discrimination task in the presence of static face, dynamic facial motion and non-face visual stimuli, and recorded single-neuron activity from the ventrolateral prefrontal cortex.

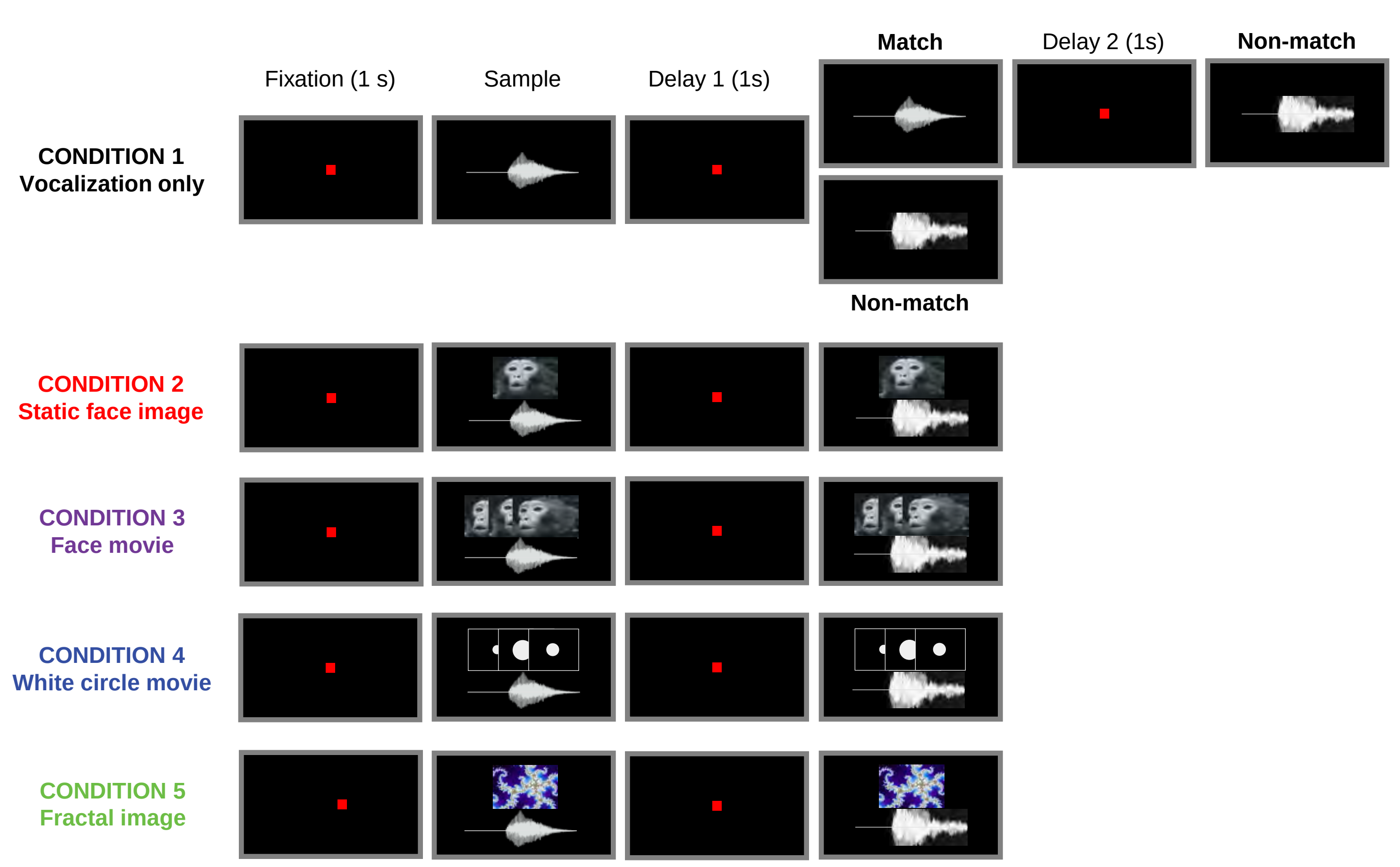
Methods

Task. A trial began when the animal fixated a point in the center of the screen. A species-specific vocalization was then presented as the SAMPLE. Additional stimuli would be presented subsequently that would match (MATCH stimulus) or not match (NON-MATCH stimulus) the sample vocalization. The subject would press a button when the NON-MATCH stimulus was presented. In some trials the second vocalization was the non-match (0-intervening trials), while in other trials the sample was repeated and the third vocalization was the non-match (1-intervening trials). Successful detection of the non-match vocalization with a button press resulted in a juice reward. A trial was aborted if fixation was broken during sound presentation or when a button press occurred before the non-match.

Visual Stimuli. There were five stimulus conditions. In condition 1, only the vocalization to be discriminated was presented. In conditions 2, 3, 4 and 5, a visual stimulus accompanied the vocalization as shown below. The correct response was entirely dependent on the successful discrimination of the vocalization which differed in the non-match period, thus the visual stimuli were irrelevant.

Training. Initially the animal was trained in the vocalization-only condition. Then, face images and fractal movies were added when the animal succeeded in over 80% of trials. Finally, face movies and white circle movies were included when the animal responded correctly in 70% of trials.

Vocalization discrimination task

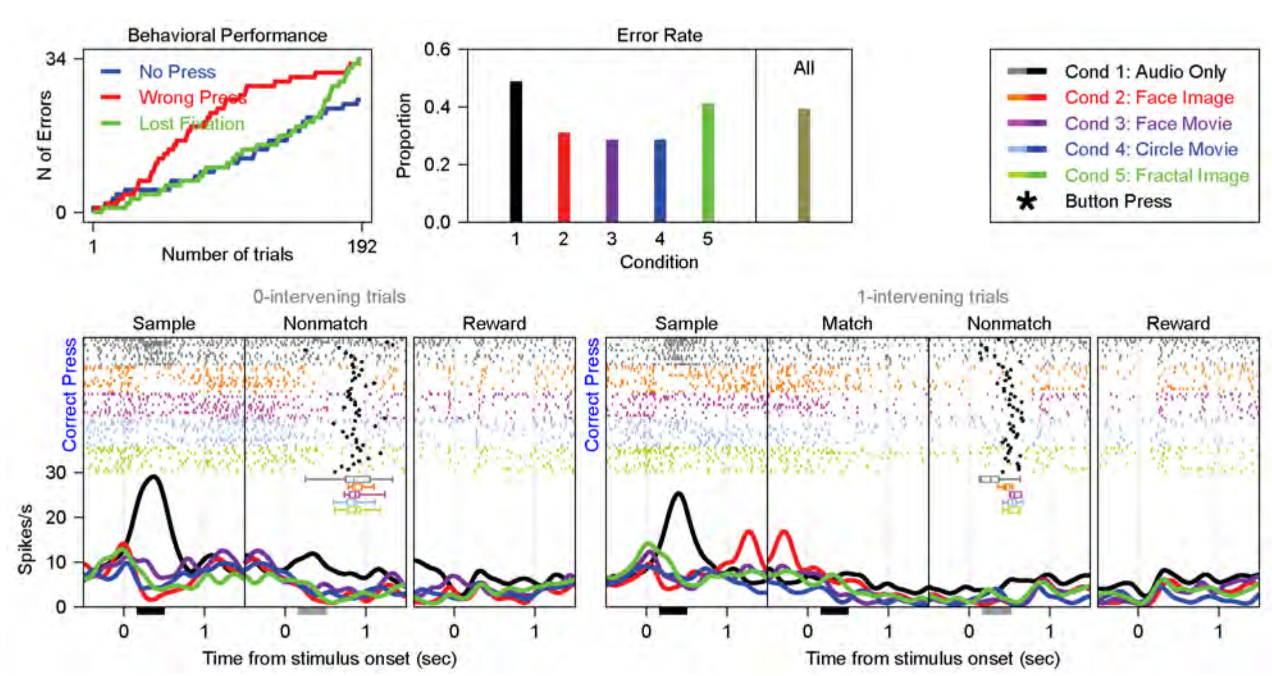


Recording. Single-neuron activity was recorded from the cylinder placed over the lateral prefrontal cortex (areas 12, 45 and 46) and was isolated using a signal processing system (Pentusa system, Tucker-Davis Technology, FL). The schematic brain at right shows the approximate location of the recording cylinder (dotted circle) in prefrontal cortex.

Results

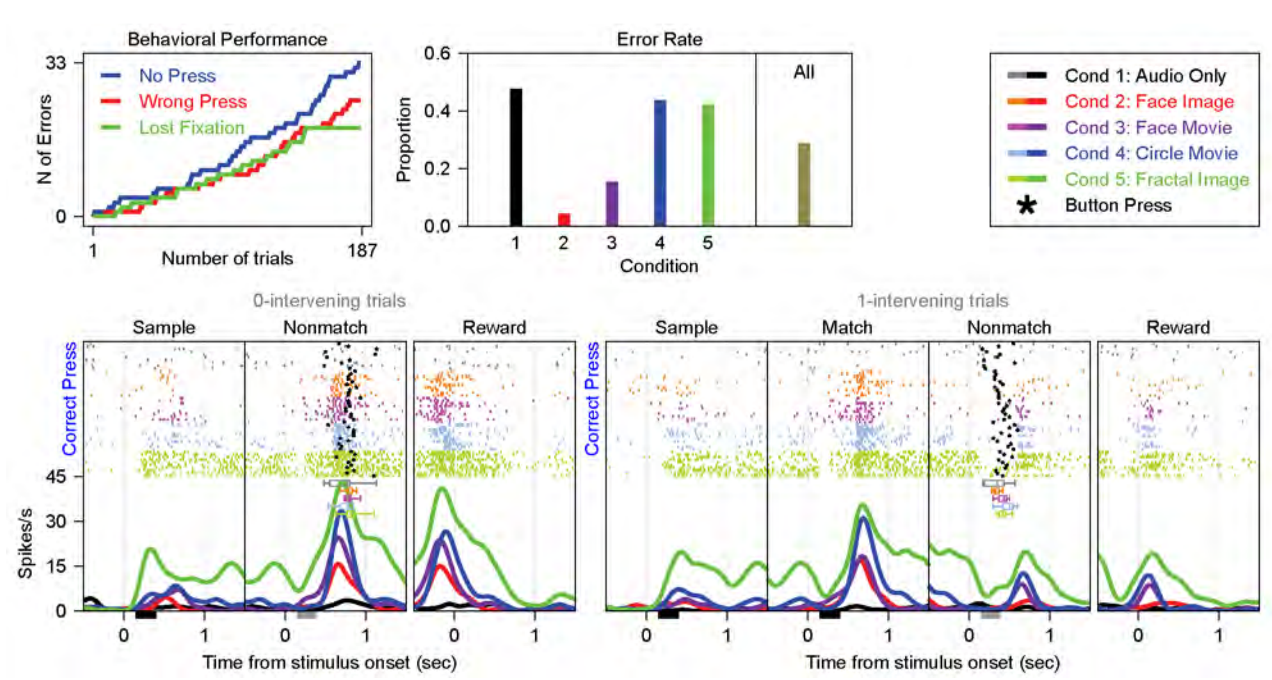
BEHAVIORAL PERFORMANCE. Subjects performed 2980 correct trials over 78 recording sessions. The average error rate was 40%. Accuracy was higher when vocalizations were presented with visual stimuli (condition 2-5), compared to the vocalization-only condition (condition1). In addition, accuracy was higher when static or motion face stimuli were presented (condition 2 & 3) than when non-face stimuli were presented (condition 4 & 5) with vocalizations. There was a small improvement in performance with motion stimuli (condition 3, facial motion & 4, expanding circle), but the effect was small compared to the effect of faces.

Fig 1. Neuron responding to the vocalization-only condition



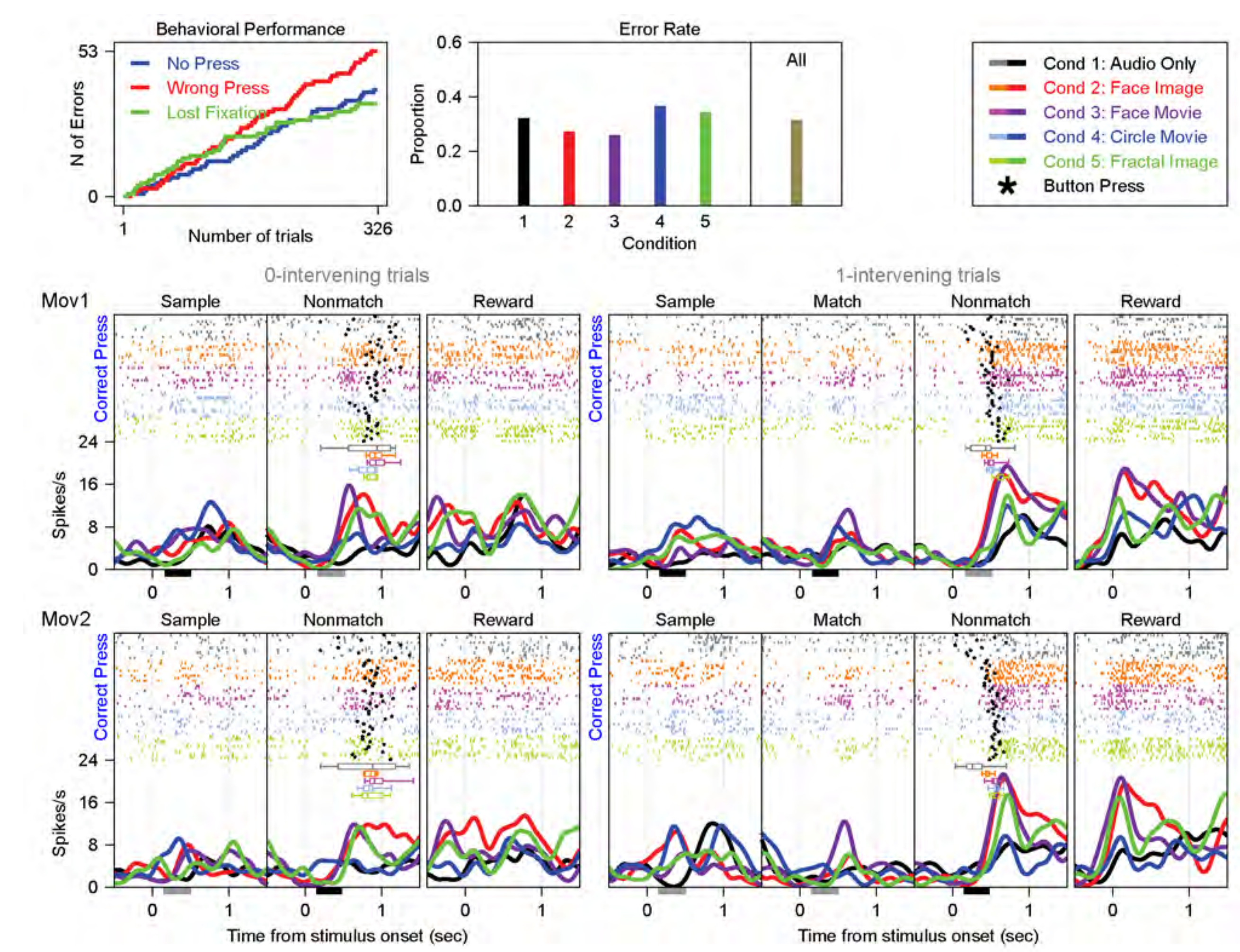
This neuron responded actively during the sample period but only when the vocalization was presented without visual stimuli. Simultaneous presentation of visual stimuli inhibited the response to the vocalization. This type of multisensory integration is known as multisensory suppression.

Fig 2. Neuron responding to visual-stimulus conditions



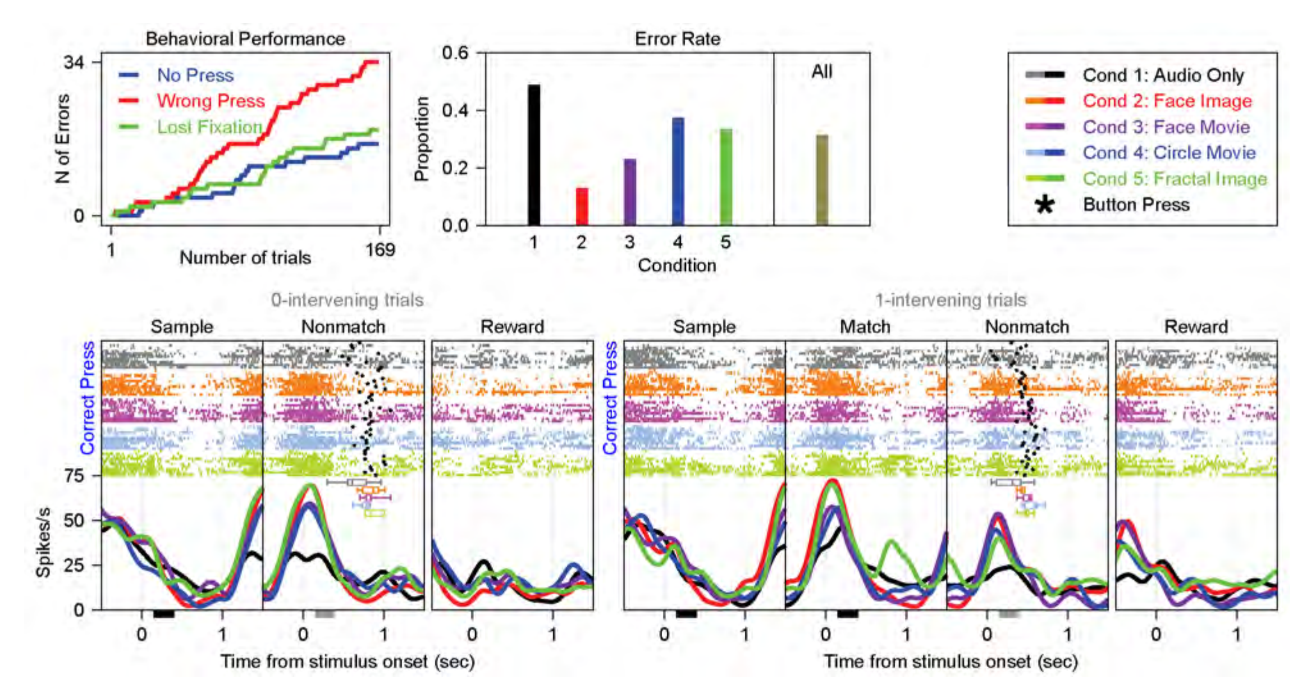
This neuron was not responsive to vocalization alone, but was active when visual stimuli were shown with the vocalization during the match and non-match periods. Note that although the fractal image elicited the highest firing rate, the performance was best with face stimuli.

Fig 3. Neuron with differential activity to faces vs non-faces



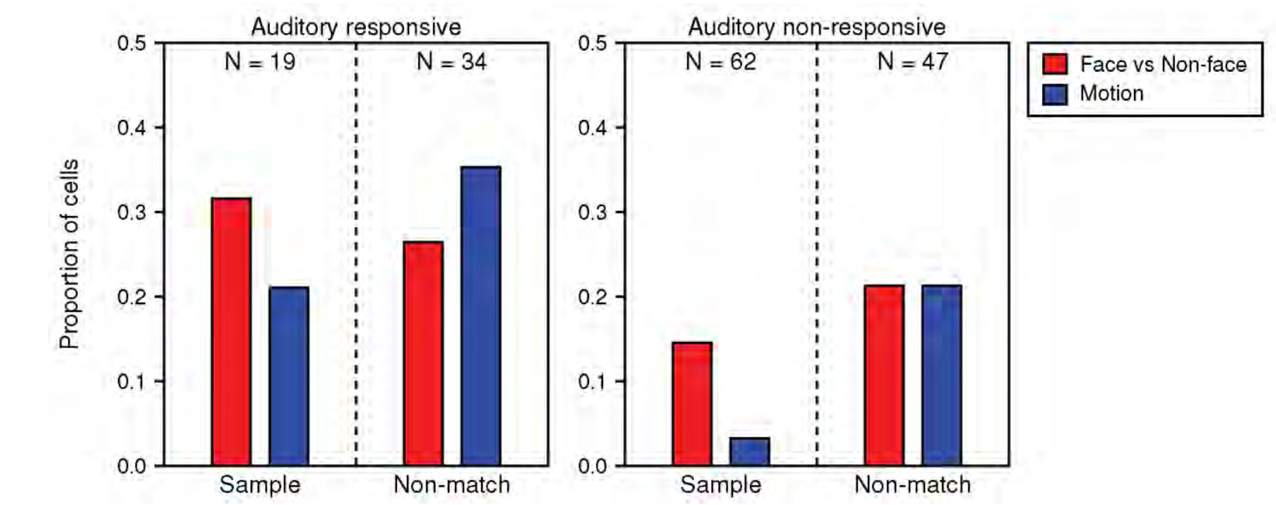
This is one example of a task related, face-selective neuron. It was responsive to all stimulus conditions, but it showed face selective enhancement during the non-match period only for the static face and face movie conditions.

Fig 4. Neuron with differential activity to motion vs static stimuli



This neuron also responded robustly to all the conditions. However, in the match and non-match epochs, the response was the highest when the vocalization was played with static images (condition 2, face image and 5, fractal image), compared to the conditions in which motion stimuli were presented. Activity of this type may occur when processing vocalizations with coherent mouth movement.

Population Result



It has been suggested that natural, dynamic facial motion stimuli which accompany vocalizations may be processed and integrated more readily than static faces presented with vocalizations or, even worse, non-face stimuli. We tested this in two steps. First, neuronal response in the VOCALIZATION ONLY condition (Condition 1) was assessed by performing a paired t-test between activity during the inter-trial interval and various task epochs. Second, the trials in which visual stimuli were presented together with vocalizations (condition 2-5) were classified into categories depending on whether the stimulus was a face or non-face (factor 1: FACE vs. NON-FACE) or whether it had a motion component (factor 2: MOTION). Then, a 2-way ANOVA was performed to assess the statistical significance of each factor (FACE and MOTION) on the task epochs Sample and Non-Match.

This analysis revealed that, during the SAMPLE period, the proportion of neurons which showed differential activity to FACE vs. NON-FACE was higher than the proportion of neurons which responded to MOTION. However, during the non-match period, when the discrimination choice must be made, more neurons showed differential activity to MOTION as opposed to static stimuli. In addition, the overall proportion of neurons which respond differently to FACE (both motion and static) vs NON-FACE stimuli was higher among the neurons responsive to the vocalizations.

Conclusions

1. Both static faces and face movies enhance the behavioral performance of vocalization discrimination more than non-face stimuli do. The effect of face was stronger than that of motion.
2. Some neurons in the lateral prefrontal cortex of non-human primates show differential activity to the visual stimuli presented with vocalizations even when the visual stimuli were not behaviorally relevant, and integrate them together.
3. Some neurons in the lateral prefrontal cortex seem to process face and facial motion stimuli differently depending on the demands of particular task epochs.
4. Neurons which are responsive to species-specific vocalizations may be more likely to be affected by accompanying face stimuli compared to non-auditory cells.

Acknowledgments

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