Behavioural Brain Research 207 (2010) 61–69

Contents lists available at ScienceDirect

Behavioural Brain Research

journal homepage: www.elsevier.com/locate/bbr

Research report

Integration of auditory and visual information in human face discrimination in pigeons

Behavioral and anatomical study

Shigeru Watanabe[a,b,*], Sayako Masuda[b]

[a] Department of Psychology, Keio University, Mita 2-15-45, Minato-Ku, Tokyo 108, Japan
[b] Centre for Advance Research on Logic and Sensibility, Keio University, Mita 2-15-45, Minato-Ku, Tokyo 108, Japan

ARTICLE INFO

Article history:
Received 30 June 2009
Received in revised form
18 September 2009
Accepted 24 September 2009
Available online 2 October 2009

Keywords:
Sensory integration
Visual discrimination
Pigeon
Social cognition
Neuroanatomy

ABSTRACT

Social stimuli are often multi-sensory. Animals commonly recognize a conspecific through visuo-auditory and olfactory sensory cues. In this study pigeons were trained to discriminate between a human face and the face of a Japanese macaque then tested with morphing images of human and macaque faces (Experiment 1). These images were also presented in conjunction with two different auditory stimuli, namely, vocalization of either a human or a macaque. Results showed that the human vocalization facilitated pigeons' recognition of human faces, but a comparable facilitation of recognition of macaque faces accompanied by a macaque vocalization was not observed. The subjects had previous exposure to human vocalizations (due to living in laboratory cages) but not to the macaque face or vocalization. This difference in experience can explain differences that such auditory stimuli have on visual discrimination in pigeons. In primates, the primary auditory area and the primary visual area connect to each other in the telencephalon. An anatomical study (Experiment 2) provided evidence that visual and auditory pathways were independent at the level of the primary sensory area in the telencephalon (entopallium and field L) suggesting an anatomical difference in visuo-auditory integration between primates and pigeons. Thus, the visual and auditory systems should be integrated in higher associative area, such as nidopallium caudo-lateralis (NCL) or in an earlier stage, such as the midbrain in the avian brain.

1. Introduction

Human beings have a well-developed ability to recognize many different human faces. Because facial features such as the eyes, nose and mouth are always positioned in a similar manner, discrimination among faces requires analysis at a relatively refined level (see [1]). Neuropsychological studies of prosopagnisia suggest a particular region of the brain (fusiform gyrus) is involved in human face perception [2], although patients showing prosopagnosia also often showed deficits in discrimination of non-face objects [3]. Brain imaging studies also suggested that the fusiform gyrus is involved in developing expertise of unfamiliar objects [4], as well as of birds and cars [5]. Thus, the identification of areas specific to human face recognition is still a matter of debate. Behavioral studies have shown that human face recognition is different from the recognition of non-face objects. Specifically, face recognition depends on configural processing of the individual features within the face, while recognition of other types of objects depends more heavily upon analyses of individual features [6,7]. Human face recognition by non-human primates has been examined using both behavioral and physiological paradigms [8–11]. It has also been investigated using birds as subjects (for example [12–15,49]).

More generally, visual recognition of conspecifics has been examined using not only humans and other primates as subjects, but also birds (chickens: [16], budgerigars: [17], pigeons: [18], Java sparrow: [19]). In each of these experiments, birds were able to discriminate individual conspecifics using visual information. Using operant conditioning, Stoddard et al. [20] successfully trained white crown sparrows to discriminate among many different conspecific song repertoires. Clearly, many bird species have the ability to discriminate individual conspecifics using visual or auditory cues.

In natural settings, effective social stimuli are often multi-sensory, thus suggesting the possibility that information from multiple sources may be integrated or that it may reveal an interactive influence on the perceiver. A typical example of an interactive effect of auditory and visual sources in human face-vocalization cognition is the McGurk effect [21]. An auditory syllable, heard in isolation as/ba/, when synchronized with a speaker who appears (visually) to be saying/ga/, is often heard as/da/. Additional

---

[*] Corresponding author at: Department of Psychology, Keio University, Mita 2-15-45, Minato-Ku, Tokyo 108, Japan. Fax: +81 3 5443 3897.
E-mail address: swat@flet.keio.ac.jp (S. Watanabe).

0166-4328/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbr.2009.09.041

support for a combined multi-sensory impact on perception comes from research indicating that speech perception is facilitated by a simultaneous visual display of a human face [22,23]. Kamachi et al. [24] also demonstrated cross-modal matching of human faces and vocalizations. These results suggest the processing of human faces by humans entails an integration of multi-sensory information. However, to date such effects mainly involve distortions of the auditory stimulus by a co-occurring visual stimulus; the converse effect, namely a perceptual modification of a visual stimulus by an auditory stimulus, has not been reported.

Animals often recognize a conspecific through visuo-auditory and olfactory sensory cues [25]. In multimodal information processing, a lack of information from one modality may be compensated through the use of other sensory channels. When conveying socially relevant information, many birds exhibit both visual and auditory signals. Recently, Partan et al. [26] showed female pigeons a videotaped playback of a male courtship display. Some playbacks contained auditory information without the video image, whereas other playbacks contained only visual information. The results showed an enhancement effect on elicited male courtship display beyond the individual effects of visual or auditory stimuli when both auditory and visual signals were presented together. Watanabe and Jian [19] trained Bengalese finches on an individual conspecific discrimination task in which still images and contact calls were simultaneously presented. The subject birds appeared to rely more on visual cues than on auditory cues when both signals arose from the same individual. However, when chimeras of visual images (e.g., a head of one bird connected with a body of another bird) were presented the subjects used the contact call for the discrimination. Thus, the dominance order of the sensory modality favors the visual modality primarily and the auditory mode secondarily; however, the secondary mode will compensate for the dominant mode when the latter does not provide sufficient information. In this case, the multi-sensory system is a kind of insurance system constructed of conspecific individual recognition when the dominant sensory information does not provide clear information. This kind of interaction has not been examined with non-conspecific social stimuli in birds. The purpose of the present experiment is to examine audio-visual interaction in discrimination of non-conspecific stimuli.

Computer-generated stimuli offer a versatile alternative to videotaped playbacks and can be used for the investigation of visual cognition in animals (see [27]). One specific form of computer-generated stimuli involves morphing, which has been used to create composite faces [28]. By changing the morphing ratio, we can produce visual stimuli along a continuum of various likenesses, namely from complete stimulus A to complete stimulus B. Intermediate stimuli have features of both A and B depending on the morphing ratio. In a previous study, following discriminative training with pigeon and starling faces in pigeons, we created morphing images of pigeons and starlings and reported the gradient of responding as a function of the morphing ratio scale [29]. In the present experiments the morphing technique was used to produce ambiguous stimuli with the aim of examining the effects of simultaneously presented auditory stimuli on the visual perception of these stimuli. Then, the effect of auditory information on discrimination of this type of ambiguous stimuli was examined.

It has been proposed that the neural basis of multi-sensory integration in primates may reside in the polymodal association area [30]. Several electrophysiological studies with monkeys have produced data that is consistent with this proposal (for example [31,32]). Moreover, the primate primary auditory area receives projections from the primary visual area [33,34]. However, it is possible that integration occurs in the primary sensory areas. Wang et al. [35] trained a monkey on an eye saccade task in which both visual and auditory stimuli were presented simultaneously. The saccade task is basically the visual oculomotor task, but the simultaneous presentation of a sound reduced saccade latencies. The V1 neurons also showed reduction in response latency suggesting that the auditory area (A1) is directly connected to the V1 visual area. Ghazanfar et al. [36] used faces and vocalizations in a study using monkeys as subjects. They found that presenting visual stimuli simultaneously with vocalization produced an enhancement of local field potential in A1. These results suggest that polymodal integration may occur at an early stage of visual information flow in monkeys. This may also be the case with birds, including pigeons and aves. In aves, for example, the tecto-fugal pathway is the main visual information projection and the entopallium is the telencephalic target of this pathway [37]. Thus, the entopallium is functionally the primary visual cortex in primates. The primary auditory cortex, which serves as analogue in birds, is field L of rose. If the visuo-auditory integration occurs at the primary sensory level as in the primate brain, there should be direct connection between the entopallium and the field L. The purpose of Experiment 2 was to conduct an anatomical examination of auditory–visual interaction between the primary sensory areas in pigeons. I injected a retrograde tracer in the entopallium and the field L separately to examine the anatomical basis of audio-visual integration in early stage of sensory processing and compare the results obtained from monkeys.

2. Experiment 1: behavioral study

Humans have a well-developed visual system; this is also true for most birds. Here, we initially trained pigeons to discriminate between human and macaque faces and then presented them morphed visual images of these faces, accompanied by auditory stimuli. Because these pigeons were familiar with humans and their vocalizations, but not with macaques, the human vocalization should generally enhance human-likeness of morphing images that have visual features of both the human and macaque faces, assuming that these birds have a visuo-auditory integrating system that operates in human face recognition. On the other hand, macaque vocalization should not have strong effects on visual discrimination of macaque faces, because the pigeons do not have cross-modal experience derived from exposure to macaques.

2.1. Methods

2.1.1. Subjects

Eight experimentally naïve homing pigeons (*Columba livia*) obtained from the Japanese Association of Racing Pigeons were used in this study. Subject birds were maintained at 80% of their free-feeding weights. Water and grit were freely available in the cages. The temperature of the animal room was maintained at 23 °C and the light cycle was maintained at 12 L:12 D.

2.1.2. Apparatus

Standard operant chambers were used (30 cm × 25 cm × 30 cm, MED). The front panel contained a rectangular transparent pecking key (10 cm × 7 cm) through which the subject could see an iMac (Power PC G4 with liquid crystal display) computer monitor. An electronic liquid shutter (UM glass, Tokyo) was placed between the key and the monitor. The distance between the key and the monitor was 15 cm. Stimuli were presented on the monitor of the iMac using power point software. A computer with a MED-SKED system controlled the experiment.

2.1.3. Stimuli

We used monochromatic images of a human face and a Japanese macaque (Macaca fuscata) face as discriminative stimuli. The human face was an image of young Japanese male of unknown age. He was completely unfamiliar for the subjects. The macaque was a young adult male, but the exact age was unknown. The human face was obtained from a stock of facial images in the Department of Psychology at Keio University and the macaque face from the Primate Research Institute at Kyoto University. The size of all stimuli was approximately 12 cm × 12 cm when displayed on the monitor screen. We employed the morphing technique using software Morpher (http://www.asahi-net.or.jp/~FX6M-FJMY/mop00j.html). Fifty points in each face (human, macaque) were selected to produce the morphing images. The mixing ratio was 0, 20, 40, 60, 80 and 100% as shown in Fig. 1. Two different modifica-