Influences of attention on auditory aftereffects following purely visual adaptation

JI HONG and THOMAS V. PAPATHOMAS*

Laboratory of Vision Research, RuCCS, and Department of Biomedical Engineering, Rutgers University, Piscataway, NJ 08854, USA

Received 18 February 2006; accepted 13 May 2006

Abstract—Recently, Kitagawa and Ichihara (2002) demonstrated that visual adaptation to an expanding or contracting disk produces a cross-modal visually-induced auditory loudness aftereffect (VALAE), which they attributed to cross-correlations of motion in three-dimensional space. Our experiments extend their results by providing evidence that attending selectively to one of two competing visual stimuli of the same saliency produces a cross-modal VALAE that favors the attended stimulus. These cross-modal attentional effects suggest the existence of integrative spatial mechanisms between vision and audition that are affected by attention.

Keywords: Cross-modal; selective attention; competing stimuli.

INTRODUCTION

There has been a renewal of interest, among psychophysicists and physiologists, in crossmodal influences in perception over the last decade or so (for reviews, see Eimer, 2004; Shimojo and Shams, 2001). This renewal has been fueled by, among other factors, advances in brain imaging techniques that have enabled researchers to investigate neural correlates of crossmodal perception (see Macaluso and Driver, 2003; Spence and Driver 2004, for reviews). These recent studies have confirmed the existence of significant crossmodal interactions, especially between auditory and visual processes in spatial perception. Thus, the classical ventriloquist illusion (Howard and Templeton, 1966), in which the location of a sound source is shifted drastically toward a simultaneously presented congruent visual stimulus, has been studied extensively under a wide variety of conditions (e.g. Alais and Burr, 2004; Bertelson, 1999; Bertelson et al., 2000; Vroomen and De Gelder, 2004).

*To whom correspondence should be addressed. E-mail: papathom@rci.rutgers.edu
Most of the early psychophysical research on crossmodal influences in the perception of visual and auditory stimuli dealt with localization of stationary crossmodal stimuli (e.g. Bertelson and Radeau, 1981; Lovelace and Anderson, 1993), but there have been increasingly more studies with moving crossmodal stimuli (e.g. Ehrenstein and Reinhardt-Rutland, 1996; Lakatos, 1995). These psychophysical experiments were accompanied by corresponding neurophysiological (e.g. Harris et al., 1980; Stein and Wallace, 1996), brain imaging (e.g. Calvert et al., 2001) and event-related potential (e.g. Fort et al., 2002) studies. One of the main findings of these neurophysiological studies was that the superior colliculus (SC) plays an important role in integrating such bimodal signals. In addition, there is evidence that two cortical sites, the anterior ectosylvian and rostral lateral suprasylvian sulci, send signals to the SC that enable it to integrate stimuli from the auditory and visual modalities (for a review, see Stein, 2005).

There have been numerous reports in which the perception of dynamic auditory events is significantly influenced by dynamic visual stimuli. Thus, the McGurk effect (McGurk and MacDonald, 1976) demonstrates that the speech sounds we perceive are influenced by the visual input from the speaker’s mouth; there are several examples where auditory motion is affected by simultaneous visual motion (Kitajima and Yamashita, 1999; Mateeff et al., 1985; Vroomen and de Gelder, 2003) or is modulated as a result of adaptation to visual motion (Ehrenstein and Reinhardt-Rutland, 1996; Kitagawa and Ichihara, 2002). There have also been reports of effects in the reverse direction, in which auditory stimuli influence the processing of visual stimuli, such as the phenomenon of audio-visual motion capture (Meyer and Wuerger, 2001). Regan and Spekreijse (1977) observed that the rate of a clicking sound affected the perceived visual flicker rate. Shams et al. (2000, 2002) took this to the extreme: a single light flash is seen as multiple flashes when it is presented simultaneously with multiple sound beeps. Sekuler et al. (1997) designed a compelling effect with a visual stimulus where two disks approach each other, then stream through, and recede away from each other; the presence of an auditory click biases the percept in favor of the two disks bouncing off, rather than streaming through, one another.

However, in most of these studies, it is the presence of a strong, unambiguous, stimulus in one modality that affects the perceptual processing of a stimulus in another modality, even in cases where the two stimuli are not present simultaneously. In particular, Kitagawa and Ichihara (2002) showed that adapting to an unambiguous purely visual stimulus, i.e. an expanding or contracting disk, produces not only a visually induced visual motion aftereffect (VVMAE), but also a cross-modal visually induced auditory loudness aftereffect (VALAE); namely, a sound of fixed intensity appears to decrease or increase over time, respectively. The most parsimonious explanation of this phenomenon, which is also supported by evidence from neuroscience (Andersen, 1997; Driver and Spence, 2000; Poirier et al., 2005), is that the neural mechanisms involved in visual motion processing interact with those involved in the processing of auditory motion. The experiments in this study