Neuroscience and Biobehavioral Reviews 61 (2016) 208–224

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

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

Review

The interactions of multisensory integration with endogenous and exogenous attention

Xiaoyu Tang[a,c], Jinglong Wu[b,c,*], Yong Shen[d,e]

[a] College of Psychology, Liaoning Normal University, 850 Huanghe Road, Shahekou District, Dalian, Liaoning, 116029, China
[b] Key Laboratory of Biomimetic Robots and System, Ministry of Education, State Key Laboratory of Intelligent Control and Decision of Complex Systems, Beijing Institute of Technology, 5 Nandajie, Zhongguancun, Haidian, Beijing 100081, China
[c] Biomedical Engineering Laboratory, Graduate School of Natural Science and Technology, Okayama University, 3-1-1 Tsushima-naka, Okayama, 700-8530, Japan
[d] Neurodegenerative Disease Research Center, School of Life Sciences, University of Science and Technology of China, CAS Key Laboratory of Brain Functions and Disease, Hefei, China
[e] Center for Advanced Therapeutic Strategies for Brain Disorders, Roskamp Institute, Sarasota, FL 34243, USA

ARTICLE INFO

Article history:
Received 22 December 2014
Received in revised form 1 November 2015
Accepted 2 November 2015
Available online 10 November 2015

Keywords:
Multisensory integration
Multisensory processing
Attention
Exogenous attention
Endogenous attention
Attentional selectivity
Multisensory search templates
Cross-modal spread of attention

ABSTRACT

Stimuli from multiple sensory organs can be integrated into a coherent representation through multiple phases of multisensory processing; this phenomenon is called multisensory integration. Multisensory integration can interact with attention. Here, we propose a framework in which attention modulates multisensory processing in both endogenous (goal-driven) and exogenous (stimulus-driven) ways. Moreover, multisensory integration exerts not only bottom-up but also top-down control over attention. Specifically, we propose the following: (1) endogenous attentional selectivity acts on multiple levels of multisensory processing to determine the extent to which simultaneous stimuli from different modalities can be integrated; (2) integrated multisensory events exert top-down control on attentional capture via multisensory search templates that are stored in the brain; (3) integrated multisensory events can capture attention efficiently, even in quite complex circumstances, due to their increased salience compared to unimodal events and can thus improve search accuracy; and (4) within a multisensory object, endogenous attention can spread from one modality to another in an exogenous manner.

Contents

* Corresponding author at: Key Laboratory of Biomimetic Robots and System, Ministry of Education, State Key Laboratory of Intelligent Control and Decision of Complex Systems, Beijing Institute of Technology, 5 Nandajie, Zhongguancun, Haidian, Beijing, 100081, China. Tel.: +81 86 251 8052; fax: +81 86 251 8266
E-mail address: wu@mech.okayama-u.ac.jp (J. Wu).

http://dx.doi.org/10.1016/j.neubiorev.2015.11.002
0149-7634/

1. Introduction

1.1. Multisensory integration

When we look for a friend in a rowdy crowd, it is easier to find our target if that person waves his/her arms and shouts loudly. To help us complete this search task more rapidly, information from different sensory modalities (i.e., visual: the waving arms; and auditory: the shout) not only interacts but also converges into a coherent and meaningful representation. These interactions and convergences between individual sensory systems have been termed multisensory integration (Lewkowicz and Ghazanfar, 2009; Talsma et al., 2010). There are two main types of behavioral outcomes of multisensory integration. The first type includes the multisensory illusion effects that have been demonstrated to illustrate the merging of information across senses, e.g., the ***ventriloquism effect***[1] (Hairston et al., 2003), the ***McGurk effect*** (McGurk and MacDonald, 1976), the ***freezing effect*** (Vroomen and de Gelder, 2000), and the ***double-flash illusion*** (Shams et al., 2000). The second type includes multisensory performance improvement effects, such as the redundant signals effect (RSE), in which responses to the simultaneous presentation of stimuli from multiple sensory systems can be faster and more accurate than responses to the same stimuli presented in isolation (Hershenson, 1962; Kinchla, 1974). In this paper, we focus on the multisensory performance improvement effects, such as the RSE, that are used to underscore the combining of information from separate modalities.

The multisensory integration research field has produced enormous gains in interest and popularity since the late 19th century (Stratton, 1897). In the last few decades, many studies have used technological advances in neuroimaging and electrophysiology to address where and when multisensory integration should be expected. Evidence of multisensory processing has been demonstrated in a number of cortical and subcortical human brain areas (see Fig. 1a). The superior colliculus (SC) is part of the midbrain and contains a large number of multisensory neurons that play an important role in the integration of information from the somatosensory, visual and auditory modalities (Fairhall and Macaluso, 2009; Meredith and Stein, 1996; Wallace et al., 1998). The superior temporal sulcus (STS), which is an association cortex, mediates multisensory benefits at the level of object recognition (Werner and Noppeney, 2010b), especially for biologically relevant stimuli from different modalities; such stimuli include speech (Senkowski et al., 2008), faces/voices (Ghazanfar et al., 2005), and real-life objects (Beauchamp et al., 2004; Werner and Noppeney, 2010a). Posterior parietal regions such as the superior parietal lobule (SPL) and intraparietal sulcus (IPS) can mediate behavioral multisensory facilitation effects (Molholm et al., 2006; Werner and Noppeney, 2010a) through anticipatory motor control (Krause et al., 2012b). The posterior parietal and premotor cortices act at guiding and controlling action in space and are also important for the integration of neural signals from different sensory modalities (Bremmer et al., 2001; Driver and Noesselt, 2008). Prefrontal cortex neurons have been found to participate in meaningful cross-modal associations (Fuster et al., 2000). For example, the ventrolateral prefrontal cortex (vlPFC) mediates multisensory facilitation of semantic categorization (Sugihara et al., 2006; Werner and Noppeney, 2010a). Moreover, integration between the senses can influence activity at some of the lowest cortical levels, e.g., the primary visual cortex (Martuzzi et al., 2007; Romei et al., 2007), primary auditory cortex (Calvert et al., 1997; Van den Brink et al., 2014), and primary somatosensory cortex (Cappe and Barone, 2005; Zhou and Fuster, 2000). These presumptive unimodal sensory areas have also been suggested to be multisensory (Ghazanfar and Schroeder, 2006).

In addition, multisensory integration has been attributed to anatomical connections between different brain areas. On the one hand, connections between sensory-related subcortical structures and the corresponding cortical areas play a role in multisensory processing. Such connections include those between the medial geniculate nucleus (MGN) and primary auditory cortex and between the lateral geniculate nucleus (LGN) and primary visual cortex (Noesselt et al., 2010; Van den Brink et al., 2014). Multisensory integration in the SC has also been shown to be mediated by cortical inputs (Bishop et al., 2012; Jiang et al., 2001). On the other hand, connections between cortical areas can mediate multisensory improvements. For example, synchronous auditory stimuli may amplify visual activations by increasing the connectivity between low-level visual and auditory areas and improve visual perception (Beer et al., 2011; Lewis and Noppeney, 2010; Romei et al., 2009).

The neural areas that are correlated with multisensory integration (especially its improvement of behavioral/perceptual outcomes) have been summarized above. Obviously, multisensory integration can occur across multiple neural levels (i.e., at subcortical levels, at the level of association cortices, and at the lowest cortical levels), which indicates that multisensory integration can be modulated by a variety of factors. Previous studies have shown that the ***intensity***, ***temporal coincidence***, and ***spatial coincidence*** [at least in some circumstances; see the review by (Spence, 2013)] of multisensory stimuli are determinants of multisensory integration (Meredith et al., 1987; Meredith and Stein, 1986a,b; Stein and Meredith, 1993; Stein et al., 1993). Although multisensory integration is typically considered an automatic process, it can be affected by ***top-down*** factors, such as attention (Talsma and Woldorff, 2005).

1.2. Endogenous and exogenous attention

Attention plays a key role in selecting relevant and ruling out irrelevant modalities, spatial locations, and task-related objects. Two mechanisms, endogenous and exogenous, are involved in this filtering process. Endogenous attention is also called voluntary or goal-driven attention and involves a more purposeful and effort-intensive ***orienting*** process (Macaluso, 2010), e.g., orienting to a red table after someone tells you that your friend is at a red table. In contrast, exogenous attention, which is also called involuntary or ***stimulus-driven*** attention, can be triggered reflexively by a salient sensory event in the external world (Hopfinger and West, 2006), e.g., the colorful clothing of your friend causes him/her to stand out.

The relationship between endogenous and exogenous attention has been extensively explored. In studies of the visual system, endogenous and exogenous attention are generally considered to be two distinct attention systems that have different behavioral effects and partially unique neural substrates (Berger et al., 2005; Chica et al., 2013; Mysore and Knudsen, 2013; Peelen et al., 2004). Unlike endogenous attention, exogenous attention does not demand cognitive resources and is less susceptible to interference (Chica and Lupiáñez, 2009). The effects that are induced

[1] Terms with the format of ***bold-italics*** have been explained in **Glossary**. See supplementary material.