Acta Psychologica 183 (2018) 43–50

Contents lists available at ScienceDirect

Acta Psychologica

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

A biphasic effect of cross-modal priming on visual shape recognition

Sze Chai Kwok[a,b,c,*], Carlo Fantoni[d], Laura Tamburini[d], Lei Wang[a], Walter Gerbino[d]

[a] Faculty of Education, Shanghai Key Laboratory of Brain Functional Genomics, Key Laboratory of Brain Functional Genomics (Ministry of Education), School of Psychology and Cognitive Science, East China Normal University, Shanghai 200062, China
[b] Shanghai Key Laboratory of Magnetic Resonance, East China Normal University, Shanghai 200062, China
[c] NYU-ECNU Institute of Brain and Cognitive Science at NYU Shanghai, Shanghai 200062, China
[d] Department of Life Sciences, Psychology Unit "Gaetano Kanizsa", University of Trieste, Italy

ARTICLE INFO

Keywords:
Attention
Cross-modal correspondence
Recognition memory
Priming

ABSTRACT

We used a cross-modal priming paradigm to evoke a biphasic effect in visual short-term memory. Participants were required to match the memorandum (a visual shape, either spiky or curvy) to a delayed probe (a shape belonging to the same category). In two-thirds of trials the sequence of shapes was accompanied by a task-irrelevant sound (either *tʒk* or *upo*, cross-modally correspondent to spiky and curvy shape categories, respectively). The biphasic effect occurred when a congruent vs. incongruent sound was presented 200 ms after the memorandum, while it did not occur when the sound was presented 200 ms before or simultaneously with it. The biphasic pattern of recognition sensitivities was revealed by an interaction between cross-modal congruency and probe delay, such that sensitivity was higher for visual shapes paired with a congruent rather than incongruent sound with a 300-ms delay, while the opposite was true with a 1300-ms delay. We suggest that this biphasic pattern of recognition sensitivities was dependent on the task-irrelevant sound activating different levels of shape processing as a function of the relative timing of sound, memorandum, and probe.

1. Introduction

Cross-modal congruency is relevant to both perception and memory (Brunetti, Indraccolo, Mastroberardino, Spence, & Santangelo, 2017; Spence, 2011). Makovac, Kwok, and Gerbino (2014) capitalized on a sound/shape cross-modal congruency phenomenon (Makovac & Gerbino, 2010) to orient participants' attention automatically towards a selected part of a two-item memorandum in a short-term visual memory task. They suggested that – after initial facilitation – an attentional shift during encoding of the memorandum, induced by a congruent audio-visual pairing, may inhibit the recognition of the cued item (but not of the non-cued item) stored in visual working memory at a critical probe delay of around 1 s. They interpreted this impaired recognition performance in valid trials (i.e., those cued by a sound congruent with the to-be-matched probe) as an inhibition-of-return-like (IOR-like) interference with the maintenance and retrieval of the encoded item.[1] Before this IOR-like interference the attention deployment to the congruent shape actually facilitated performance at a very short probe delay. The occurrence of such a biphasic pattern of recognition performance, consisting of early facilitation (150-ms probe delay) followed by late deterioration (1150-ms probe delay), actually raises the following question, which motivated the present experiment: Can a similar biphasic effect of cross-modal congruency occur independently of the spatial orienting of attention, as when the to-be-matched memorandum and probe shapes are displayed in succession at the focus of attention?

This question is challenging for several reasons. Classically, IOR refers to the loss of effectiveness of a target presented in a previously cued, compared to an uncued, location (Dukewich & Klein, 2015; Posner & Cohen, 1984). Inhibition of orienting to previously attended locations has been proposed to facilitate foraging for novel information to aid visual search or more efficiently explore the external environment (Klein, 2000). Phenomena similar to IOR have been linked with working memory (Belopolsky, Schreij, & Theeuwes, 2010; Theeuwes, Van der Stigchel, & Olivers, 2006) and known as memorial IOR (Dukewich & Klein, 2015). For example, Johnson et al. (2013) applied a memory refreshing paradigm to create reflective attention in memory that is analogous to perceptual selective attention. By instructing participants to think back to and foreground an active mental representation of a just-presented stimulus, they found that responses were significantly slower for refreshed than unrefreshed items. This suggests that reflective attention can inhibit items represented in memory, paralleling IOR effects of perceptual attention.

* Corresponding author at: Room 269, Geography Building, 3663 Zhongshan Road North, Shanghai 200062, China.
E-mail address: sze-chai.kwok@st-hughs.oxon.org (S.C. Kwok).

[1] We acknowledge that the IOR label should be adopted cautiously (Berlucchi, 2006). The biphasic pattern obtained by Makovac et al. (2014) might depend on mechanisms different from IOR.

https://doi.org/10.1016/j.actpsy.2017.12.013
Received 24 May 2017; Received in revised form 27 December 2017; Accepted 31 December 2017
Available online 08 January 2018
0001-6918/

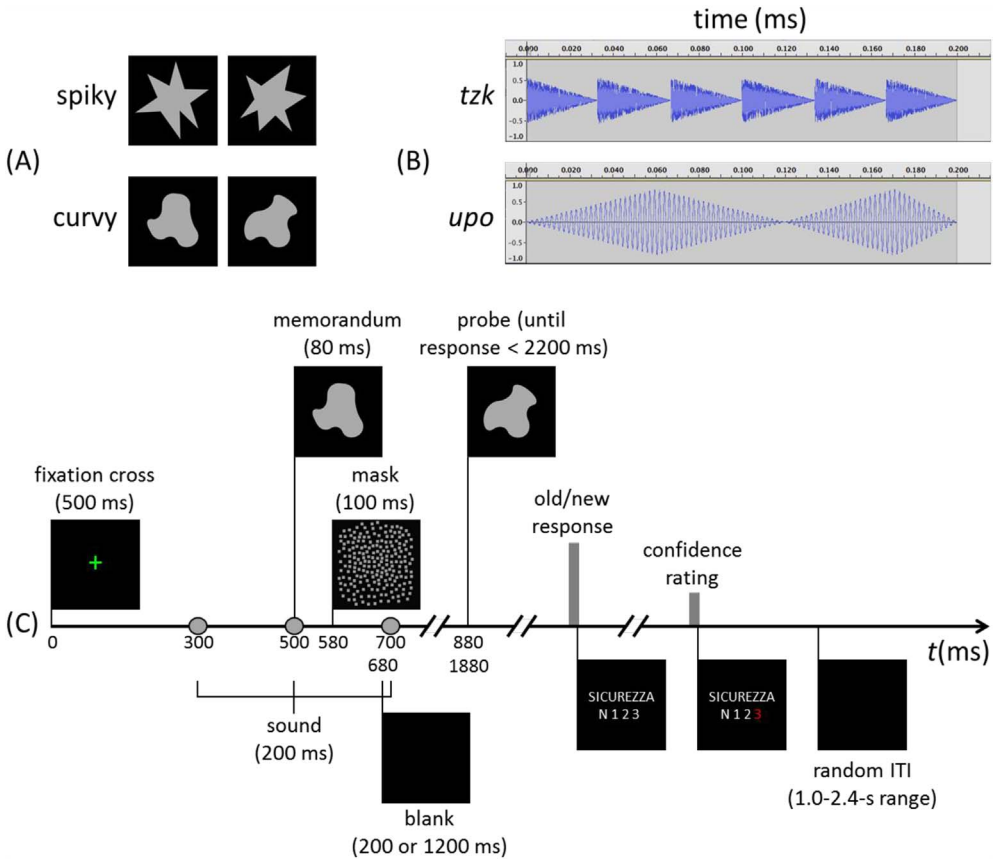

Fig. 1. (A) Examples of to-be-matched visual shapes: two spiky and two curvy (see Supplementary material for the entire stimulus set including 10 spiky and 10 curvy shapes). (B) Waveforms of *tzk* and *upo* sounds (according to Audacity® convention). (C) Trial temporal structure: this specific example illustrates the subset including negative trials with curvy shapes (probe physically different from the memorandum). Depending on sound (*upo*, *tzk*, or absent), the condition was labelled as either congruent, incongruent, or baseline. Depending on after-mask blank duration (either 200 or 1200 ms), the probe was presented after 880 or 1880 ms from the beginning of the trial, with a delay of either 300 or 1300 ms from the offset of the memorandum. The three levels of $SOA_{s\text{-}m}$ are displayed as three grey dots along the time line.

The specific instance of cross-modal congruency employed in Makovac et al. (2014) is a kind of synaesthetic correspondence between putatively non-redundant stimulus attributes or dimensions that happens to be shared by many people (Spence, 2011). Stimuli that are synaesthetically congruent are more likely to be bound together, a notion that has been referred to as the "unity effect" (Vatakis, Ghazanfar, & Spence, 2008). In the present context, we build upon the fact that this correspondence between inputs can make sounds and visual shapes perceptually similar (Köhler, 1929), speed up cross-modal event detection (Makovac & Gerbino, 2010), modulate attention (Chiou & Rich, 2012), improve recognition (Murray et al., 2004), and biphasically influence recognition performance (Makovac et al., 2014). Similar effects have also been found when the correspondence can be established between visual and motor inputs (Fantoni & Gerbino, 2014; Fantoni, Rigutti, & Gerbino, 2016). While it is established that cross-modal congruency constrains the binding of stimulus features in perception (Gerbino & Fantoni, 2016; Spence, 2011), its effects on memory representation and post-perceptual response-selection have been identified only recently (Brunetti et al., 2017).

In the present study we tested whether a biphasic effect of cross-modal congruency can occur independently of the spatial orienting of attention. Makovac et al. (2014) considered this possibility, admitting that the cueing paradigm used in their experiment could elicit either a spatial-based or a feature-based attentional shift (or both). We thus designed a sequential matching task in which a target shape was followed by a comparison shape belonging to the same category (either spiky or curvy), to be matched according to their physical identity. To be consistent with the experiment in Makovac et al. (2014), the target and comparison shapes were called "memorandum" and "probe", respectively, though in the present experiment the memorandum included only one item. In two-thirds of trials the presentation of the sequence of two shapes was accompanied by a synaesthetically congruent (*upo* → curvy; *tzk* → spiky) or incongruent (*upo* → spiky; *tzk* → curvy) sound. Such a procedure is commonly considered as a priming rather than cueing paradigm, the latter being typically associated with IOR when the delay between the cue and the imperative stimulus is sufficiently long (Dukewich & Klein, 2015).

The difficulty of the physical identity match required by our task might differ for spiky and curvy shapes, independently of the synaesthetic correspondence with *upo* and *tzk* sounds. The two sets of spiky and curvy shapes were designed to be clearly discriminable at a categorical level, to maximize the congruency/incongruency with the two sounds. This was obtained by contrasting both contour trajectories (abrupt discontinuities for spiky shapes vs. smooth changes of curvature for curvy shapes) and skeletal structures (six-pointed stars vs. three-lobed blobs). As a working hypothesis, we assumed that the expected cross-modal congruency effect was not substantially contaminated by a possible recognition superiority of either spiky or curvy shapes in the successive matching paradigm. In principle, such a superiority might emerge during encoding and/or matching, and depend on the relative complexity of closed contours (Redies, Brachmann, & Wagemans, 2017) as well as on the aesthetic preference for angular vs. curvilinear abstract shapes (Palumbo & Bertamini, 2016; Palumbo, Ruta, & Bertamini, 2015). Since we did not measure the within-category discriminability of spiky and curvy shapes before the experiment, we had no specific expectations about the interplay of complexity at the level of contour trajectory and skeletal structure, as a predictor of the overall difficulty of successive matching based on physical identity.

The expected biphasic effect of a cross-modally congruent categorical prime would consist in better performance (higher recognition sensitivity) with a congruent than incongruent sound after a short retention delay (300 ms), followed by better performance with an incongruent than congruent sound after a long retention delay (1300 ms). This interaction between congruency and delay would suggest that the biphasic effect found by Makovac et al. (2014) using a cueing paradigm might at least partially originate from a shift of feature-based attention,