Brain and Cognition 90 (2014) 19–31

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

Visual enhancement of auditory beat perception across auditory interference levels

Yi-Huang Su *

Department of Movement Science, Faculty of Sport and Health Science, Technical University of Munich, Germany

ARTICLE INFO

Article history:
Accepted 8 May 2014
Available online 6 June 2014

Keywords:
Rhythm perception
Multisensory perception
Visual beat
Point-light figure
Inverse effectiveness
Sensorimotor synchronization

ABSTRACT

This study dealt with audiovisual rhythm perception involving an observed movement. Two experiments investigated whether a visual beat conveyed by a bouncing human point-light figure facilitated beat perception of concurrent auditory rhythms, and whether this enhancement followed a profile of multisensory integration. In Experiment 1, participants listened to three repetitions of a metrically simple rhythm and detected a perturbation in the third repetition. The rhythm was presented alone or with a visual beat in phase to it. Both conditions were presented with or without an auditory interference sequence at four increasing tempi, which served to progressively weaken the beat of the auditory rhythm. In Experiment 2, participants tapped to a regular auditory beat in the same combinations of visual beat and auditory interference. Results showed that the visual beat improved the perception of (Experiment 1) and the synchronization to (Experiment 2) the auditory rhythms. Moreover, in both experiments, visual enhancement was greater when the performance in the unisensory (auditory) conditions was poorer, consistent with the principle of inverse effectiveness. The relative multisensory gain increased as auditory performance deteriorated, except in one intermediate level. Together these results demonstrate that rhythmic visual movement aids auditory rhythm perception, which may be subserved by a perceptually integrated audiovisual beat that couples the internal motor system.

1. Introduction

"Rhythm" typically refers to temporally organized patterns found in auditory stimuli such as music and speech (Large, 2008). However, humans' experiences of rhythms are often multimodal, involving visual information: e.g., we observe a musician's movement as he or she plays an instrument, we watch a dancer moving to the music, we hear the footsteps of a person while seeing the approaching gaits, or we read the lip movement of someone speaking in front of us. These scenarios exemplify how auditory rhythms are closely and pervasively connected to rhythmic human movements that either give rise to, or are coordinated with, such auditory events (Merker, Madison, & Eckerdal, 2009). The effect of movements communicated visually along with the auditory rhythms has been well examined in speech: Rhythmic facial movements, such as lip movements coupled to the 3–8 Hz rhythm of syllable production, are known to facilitate auditory speech perception (Chandrasekaran, Trubanova, Stillittano, Caplier, & Ghazanfar, 2009; Ghazanfar, 2013; Ghazanfar, Chandrasekaran, & Morrill, 2010). Although both speech and music are unique means of human communication, multisensory perception in musical rhythms remains poorly understood. For example, does the observation of a dancer's movement reinforce the perceived rhythm of the music? The present research was concerned with multisensory perception in this context. Of particular interest was whether the perception of concurrent auditory and visual rhythms – with the latter being conveyed by an observed movement coordinated with the auditory rhythm – followed similar principles to those established in multisensory integration.

Auditory musical rhythms often give rise to the perception of a regular beat, to which listeners would naturally synchronize through various overt or covert movements (Burger, Thompson, Luck, Saarikallio, & Toiviainen, 2013; Drake, Jones, & Baruch, 2000; Merker et al., 2009; Su & Pöppel, 2012). Besides perpetuating the listeners to move along, the presence of a beat facilitates rhythm perception (Grahn, 2012; Grahn & Brett, 2007; Grube & Griffiths, 2009; Povel & Essens, 1985) and motor synchronization to the rhythms (Patel, Iversen, Chen, & Repp, 2005; Repp, Iversen, & Patel, 2008). It is commonly found that the capacity for rhythm and beat perception differs between the two major senses, audition

* Address: Department of Movement Science, Faculty of Sport and Health Science, Technical University of Munich, Georg-Brauchle-Ring 62, 80992 Munich, Germany.
E-mail address: yihuang.su@tum.de

http://dx.doi.org/10.1016/j.bandc.2014.05.003
0278-2626/

and vision. When the same task was presented through comparable auditory (typically successive tones) and visual stimuli (until recently, repetitive flashes), performance in the former well surpasses that in the latter with regard to rhythmic interval timing (Grondin & McAuley, 2009), beat perception (Grahn, 2012; Grahn, Henry, & McAuley, 2011; McAuley & Henry, 2010), and sensorimotor synchronization (Jäncke, 2000; Kato & Konishi, 2006; Lorås, Sigmundsson, Talcott, Öhberg, & Stensdotter, 2012; Patel et al., 2005; Repp, 2003; see Section 1.4.2 in Repp & Su, 2013). Similarly, when pitting concurrent auditory and visual rhythms in a target–distractor paradigm (Guttman, Gilroy, & Blake, 2005; Repp & Penel, 2002; Repp & Penel, 2004), an auditory distractor affects the visual task considerably, whereas a visual distractor has minimal influence on the auditory task. The modality differences are also reflected in the relevant (sub)-cortical motor areas underlying the performance (Grahn et al., 2011; Section 4.2.6 in Repp & Su, 2013).

As rhythm tasks essentially tap onto temporal processing of sensory stimuli, the frequently observed auditory advantage seems to be explained by the *modality appropriateness hypothesis* (Welch & Warren, 1980). This hypothesis postulates that audition and vision are better at temporal and spatial tasks, respectively, and that the more competent modality for the task dominates. However, the idea of a general visual inferiority in the rhythm domain has recently been challenged, as there is growing evidence of the rhythmic capacity of dynamic visual stimuli. For example, apparent motion of an object improves visual rhythm perception (Brandon & Saffran, 2011; Grahn, 2012) and synchronization (Hove, Fairhurst, Kotz, & Keller, 2013a; Hove, Iversen, Zhang, & Repp, 2013b; Hove, Spivey, & Krumhansl, 2010), as compared to repetitive, stationary flashes that were adopted in the earlier studies. Hove et al. (2013b) even found that a bouncing ball as a visual distractor has a stronger effect than an auditory distractor for visually trained participants (video gamers). Furthermore, similar activations in putamen – an area strongly implicated in beat perception (Grahn & Brett, 2007; Grahn et al., 2011) – have been found for synchronization with a moving visual stimulus (a bar moving in space) and with an auditory metronome (Hove et al., 2013a). It thus appears that rich motion information in the visual stimuli is one key to optimizing visual rhythm processing. This may be because (periodic) visual motions, like auditory rhythms, are capable of coupling humans' internal motor system that supports rhythm and beat perception (Section 4.1.2 in Repp & Su, 2013).

Inspired by these findings, and the link between rhythm perception and human movement (Phillips-Silver & Trainor, 2007; Su & Pöppel, 2012), a recent study examined the effect of a visual stimulus whose motion profile was extracted from a natural human movement (Su, 2014b). In that study, a periodic visual stimulus was presented as a continuously bouncing human point-light figure (PLF, Johansson, 1973, see Visual stimuli in Section 2.1.2), which had been generated by recording a real human bouncing regularly. The visual beat was perceived around each bounce (maximal knee flexion) of the continuous movement, mirroring how humans would move to a musical beat (Phillips-Silver & Trainor, 2007; see also Toiviainen, Luck, & Thompson, 2010, for converging evidence of a musical pulse/beat being embodied in humans' vertical body movements). It was found that such a visual beat, when combined with an auditory beat (i.e., yielding a bimodal beat), modulated beat perception of complex auditory rhythms, and the phase of the visual beat had more influence than that of the concurrent auditory beat. In that study, however, the effect of a unimodal visual beat was less clear because the employed auditory rhythms had complex temporal structures (the 'metric complex' patterns as used in Grahn, 2012), making the imposed beat less salient. The question thus remains as to whether the given visual beat alone can influence the perception of less complex auditory rhythms.

Following this finding, the present study investigated the effect of the same visual stimulus, a bouncing PLF, on the perception of metrically simple auditory rhythms. Specifically, and as has not been shown before, it examined whether the perception of concurrent audiovisual rhythms followed one of the principles that are often found to characterize multisensory perception: the *principle of inverse effectiveness* (PoIE). The PoIE was initially established in the firing rates of cats' superior colliculus (SC) neurons that respond to both unisensory and multisensory stimuli (Alvarado, Vaughan, Stanford, & Stein, 2007; Meredith & Stein, 1983; Stein, Stanford, Ramachandran, Perrault, & Rowland, 2009). It describes that multisensory information is combined in such a manner, that the response gain associated with multisensory cues is greater when the (most effective) unisensory stimulus strength – and accordingly the unisensory response – is weaker. For example, the weaker the response of SC neurons to a unisensory cue (e.g., the sight of an object), the more beneficial a multisensory cue (e.g., adding sound to the sight of an object) will be in eliciting neuronal response and its behavioral consequence (e.g., speeding up orienting to the object. See Stein & Stanford, 2008). This pattern has since been shown in various human behavioral measures (e.g., RT or detection) as well as cortical activations (e.g., in the superior temporal sulcus), in the perception of both simple events (e.g., flash and beep) and complex multisensory stimuli, such as speech, objects, motion, or body gestures (Hecht, Reiner, & Karni, 2008; Jessen, Obleser, & Kotz, 2012; Rach & Diederich, 2006; Saldern & Noppeney, 2013; Senkowski, Saint-Amour, Höfle, & Foxe, 2011; Stevenson & James, 2009; Stevenson et al., 2012; Werner & Noppeney, 2010). Considering the possibly overlapping mechanisms underlying speech and music cognition (Patel, 2012; Patel, 2014), that the perception of both relies on their rhythms in a similar manner (Cason & Schön, 2012; Hausen, Torppa, Salmela, Vainio, & Särkämö, 2013; Rothermich, Schmidt-Kassow, & Kotz, 2012), and that both can be conveyed through multisensory information, it may be hypothesized that multisensory perception of musical rhythms would also follow the PoIE. The present study thus asked a novel question: Does a visual beat assist auditory beat perception, such that the benefit of a visual beat increases as the auditory beat is made less distinct by a source of rhythmic interference?

This hypothesis was tested in two experiments, borrowing the paradigm from audiovisual speech perception, in which auditory word recognition with and without a visual cue (a talking face) was measured across increasing levels of auditory noise (Ross, Saint-Amour, Leavitt, Javitt, & Foxe, 2007). Here, the perception of (Experiment 1) and the synchronization to (Experiment 2) beat-based auditory rhythms with and without a concurrent visual beat were measured across increasing levels of auditory interference. In Experiment 1, the auditory rhythms were metrically simple, and the beat arises readily in the perception of these rhythms by means of temporal (grouping) accents (Grahn, 2012; Povel & Essens, 1985). A bouncing PLF as a visual beat, if present, accompanied the rhythms synchronously in the same meter. The auditory interference was *not* of the noise sort that masked the signal of the auditory rhythms. Rather, it presented interference in the rhythm domain, implemented as an isochronous sequence (of a different timber from that of the rhythms) whose period differed from the beat period of the auditory rhythms. The periods of the rhythm and of the interference were related by an integer ratio other than N:1, yielding a polyrhythm that weakened the beat saliency of the auditory rhythm (Poudrier & Repp, 2013). The level of interference was increased by elevating the tempo of the interference sequence across conditions, while keeping the tempo of the auditory rhythms constant, thereby increasing the auditory