Acta Psychologica 152 (2014) 67–74

Contents lists available at ScienceDirect

Acta Psychologica

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

Know thy sound: Perceiving self and others in musical contexts

Vassilis Sevdalis [a,*], Peter E. Keller [b]

[a] Institute of Psychology, German Sport University Cologne, Germany
[b] The MARCS Institute, University of Western Sydney, Sydney, NSW, Australia

ARTICLE INFO

Article history:
Received 16 December 2013
Received in revised form 7 April 2014
Accepted 7 July 2014
Available online 13 August 2014

PsycINFO classification:
2300 (Human Experimental Psychology)
2500 (Physiological Psychology & Neuroscience)
2200 (Psychometrics & Statistics & Methodology)

Keywords:
Perception
Action
Self-recognition
Music
Cue integration
Embodied cognition

ABSTRACT

This review article provides a summary of the findings from empirical studies that investigated recognition of an action's agent by using music and/or other auditory information. Embodied cognition accounts ground higher cognitive functions in lower level sensorimotor functioning. Action simulation, the recruitment of an observer's motor system and its neural substrates when observing actions, has been proposed to be particularly potent for actions that are self-produced. This review examines evidence for such claims from the music domain. It covers studies in which trained or untrained individuals generated and/or perceived (musical) sounds, and were subsequently asked to identify who was the author of the sounds (e.g., the self or another individual) in immediate (online) or delayed (offline) research designs. The review is structured according to the complexity of auditory–motor information available and includes sections on: 1) simple auditory information (e.g., clapping, piano, drum sounds), 2) complex instrumental sound sequences (e.g., piano/organ performances), and 3) musical information embedded within audiovisual performance contexts, when action sequences are both viewed as movements and/or listened to in synchrony with sounds (e.g., conductors' gestures, dance). This work has proven to be informative in unraveling the links between perceptual–motor processes, supporting embodied accounts of human cognition that address action observation. The reported findings are examined in relation to cues that contribute to agency judgments, and their implications for research concerning action understanding and applied musical practice.

Contents

1. Introduction

Human beings experience a vast amount of auditory information in their everyday environments, such as the growl of thunder or the horns of cars, the voices of colleagues or the snoring of neighbors. Fortunately, certain sounds take the form of music. Music has the potential to be pleasurable in an aesthetic sense, but it can also be used as a means to investigate human capabilities related to it: these may include the production of sounds and their perception — the latter related to processes of identification and recognition of physical sound properties. Sometimes, the generator and receiver of the sound are one and the same person. This review article examines relationships between sounds generated by individuals' actions and the recognition of these sounds and the individuals who produce them via listening.

Certain performing arts, such as music and dance, include a prominent auditory component in the form of sounds generated by, or produced as an accompaniment to, the performers' actions. In these cases, being able

* Corresponding author at: German Sport University Cologne, Institute of Psychology, Am Sportpark Müngersdorf 6, 50933 Cologne, Germany. Tel.: +49 221 4982 5721; fax: +49 221 4982 8320.
E-mail address: v.sevdalis@dshs-koeln.de (V. Sevdalis).

http://dx.doi.org/10.1016/j.actpsy.2014.07.002

to articulate a distinctive performance style is considered an asset and a quest pursued through long-term practice. Indeed, the actions of skilled and novice musicians and dancers have distinctive individual characteristics, that stem from anatomical constraints and different learning histories (e.g., Repp, 1992, 1995; Sevdalis & Keller, 2012). In the music domain, differences in individual variation of performance execution can occur both within an individual and between individuals. Discrepancies and commonalities in performance displays are nevertheless perceptually detectable, based on cues such as tempo, expressive timing, and dynamics (e.g., loudness). In perceptual experiments, for example, this has been demonstrated in jazz musicians accurately detecting whether the same piano melodies are improvised or imitated (Engel & Keller, 2011), in subjective judgments of similarity between different performances of the same piece (Timmers, 2005), and in aesthetic judgments of averaged individual music performances (Repp, 1997). Thus, accuracy in perceptual tasks that require identifying (di)similarities between actions is attainable when the sole information source is the sound alone.

From a theoretical point of view, the above examples showcase the capacity of body movements and their effects to shape cognitive operations, a fundamental premise of embodied cognition approaches (e.g., Grafton, 2009; Wilson, 2002). Such approaches converge on the assumption that high-level cognitive functions are grounded in low-level sensorimotor functions. According to this rationale, the importance of actions becomes particularly apparent if one considers their potential to be the means of enacting upon the environment (Herwig, Beisert, & Prinz, 2013). Essentially, performing and perceiving actions constitute an individual's means to interact with the environment and with other individuals. The functions of actions, thus, go beyond their motor components, and extend to cognitive and affective ones. Indeed, at both neurophysiological and behavioral levels, evidence is mounting that the coupling between action perception and action execution is boosted with increases in the degree to which an individual has physical experience in performing an action (e.g., Schubert & Semin, 2009; Sevdalis & Keller, 2011b).

Although research in action understanding has traditionally focused on the visual modality (but see Shams & Kim, 2010, for a review on the modulation of vision by auditory information), recent work has highlighted the importance of the auditory modality (for a review, see Aglioti & Pazzaglia, 2010). The significance of audition becomes particularly evident if one considers the diversity of activities that include primarily or solely the auditory channel, such as speaking, singing, and instrumental music performance. Evidence suggests that action-related sounds activate premotor areas in the human brain (Gazzola, Aziz-Zadeh, & Keysers, 2006), and that training on a musical instrument induces differences in somatosensory, auditory, and motor cortical brain functions and structures (e.g., Bangert et al., 2006; D'Ausilio, Altenmuller, Olivetti Belardinelli, & Lotze, 2006; Lahav, Saltzman, & Schlaug, 2007; Münte, Altenmüller, & Jäncke, 2002). Similar findings regarding the effects of action-related sounds were obtained when sounds of actions were presented to congenitally blind individuals (Ricciardi et al., 2009). Taken together, these results suggest that auditory–motor mappings are established in the brain and support the mutual influences between auditory and motor processes.

One specific class of actions that are particularly well suited for investigating these auditory–motor brain mappings and their behavioral counterparts are those produced by oneself. The perception of self-produced actions and their sounds benefits from the fact that, in this case, the observer's action system has specialized proprioceptive knowledge that is based on direct motor experience (Wilson & Knoblich, 2005). Agency, or being the agent of an action, refers to the feeling of being in control of one's actions and their effects (Pacherie, 2012; Repp & Knoblich, 2007). In the case of perceiving one's own sounds, the auditory–motor mappings have the potential to share a common code, that is, to become matched on the level of common auditory–motor representations (cf., Herwig et al., 2013). This matching of sensory and motor features of actions allows an individual to use his or her motor system to simulate an observed action, which can then be used to determine authorship (e.g., self or others) based on sensorimotor discrepancies and similarities between the simulation and the action (cf. Hommel, Müsseler, Aschersleben, & Prinz, 2001; Prinz, 1990). Such self-recognition capacities appear early in human life, between the first and second years (for reviews, see Butterworth, 1992; Rochat, 1998), are shared with other species (for a review, see Byrne & Bates, 2010), and are considered to be a constituent part of social cognition (Decety & Sommerville, 2003). However, the primary modality employed in self-recognition experiments remains the visual one (for a review, see Suddendorf & Butler, 2013). In spite of cross-modal and unimodal designs having been employed to assess self-recognition based on auditory information in an increasing number of studies (e.g., seeing and hearing one's name, Platek, Thomson, & Gallup, 2004; listening to auditory signals generated by one's footsteps, Menzer et al., 2010), a comprehensive account that deals with music-related actions is still wanting.

What makes sounds – and especially music – significant? Sounds have the potential to cover a broad coverage of environmental events and can intrinsically occur in synchrony with actions. Although vision may often dominate the human sensorimotor landscape (Colavita, 1974; Posner, Nissen, & Klein, 1976), visual information is less important in activities when auditory information is the primary means of expression, such as music-related ones (e.g., instrument learning and performance). Music is an ancient and culturally widespread activity, naturally present in most people's lives, and practiced by individuals with varying levels of expertise. These characteristics render musical sounds ecologically valid stimulus materials that can be readily used in experimental contexts. In music performance, for instance, the production of complex sequences is typically accompanied by receiving instant auditory feedback for the actions one performs: this auditory information can be experimentally manipulated to test how it affects performance execution (e.g., Pfordresher, Keller, Koch, Palmer, & Yildirim, 2011). Audition has very accurate temporal resolution: for example, the threshold for auditory temporal order judgments is around 20 ms (Hirsch, 1959; Hirsh & Watson, 1996) and the threshold for the detection of auditory onset asynchronies can be as low as 2 ms (Zera & Green, 1993). Audition is often considered a more 'accurate' sense than vision in certain situations such as temporal processing or synchronization (Arrighi, Alais, & Burr, 2006; Repp & Penel, 2002, 2004), and is more developed than vision before and at birth (Robinson & Sloutsky, 2004). Thus, sounds possess unique qualities, ranging from low-level physical properties up to high-level social information.

However, musical sounds have an inherent ambiguity attached to them (McGuiness & Overy, 2011). An interesting characteristic of musical listening is that it can create an auditory landscape whose properties are fluid in nature (i.e., they change each moment as the music unfolds) and are to some degree unpredictable (i.e., due to this continuous temporal evolution). The auditory system has lower spatial resolution than the visual system: when visually observing objects or events, both spatial and temporal dimensions can be employed for perceptual decision-making. The embodied nature of music perception and production (in terms of auditory–motor overlap) can pose challenges: essentially, identifying the properties of auditory recordings and understanding the agent's communicative intentions entails simulating the properties of actions that generated them (Keller, 2008). As a consequence, it can be more difficult to attribute one's own agency to an auditory signal, in comparison to a visual one (Sevdalis & Keller, 2010).

Musical activities, such as coordinating with a co-performer and predicting upcoming events, require monitoring one's own and others' actions (Keller, 2014; Keller, Novembre, & Hove, in press), and rely on knowing one's own and others' actions and their effects (agency). Presumably, if there is an efficient self–other distinction of sounds, then self- and other-awareness increases, and, thus, self and other sounds can merge into a coherent Gestalt during performance. Considering the universality of musical behavior, multiple factors influence music