Title
Motor simulation theories of musical beat perception

Permalink
https://escholarship.org/uc/item/6g51j0ws

Authors
Ross, Jessica
Iversen, John Rehner
Balasubramaniam, Ramesh

Publication Date
2016

Peer reviewed

REVIEW

Motor simulation theories of musical beat perception

Jessica M. Ross[a], John R. Iversen[b] and Ramesh Balasubramaniam [iD][a]

[a]Cognitive and Information Sciences, University of California, Merced, CA, USA; [b]Swartz Center for Computational Neuroscience, Institute for Neural Computation, University of California, San Diego, CA, USA

ABSTRACT

There is growing interest in whether the motor system plays an essential role in rhythm perception. The motor system is active during the perception of rhythms, but is such motor activity merely a sign of unexecuted motor planning, or does it play a causal role in shaping the perception of rhythm? We present evidence for a causal role of motor planning and simulation, and review theories of internal simulation for beat-based timing prediction. Brain stimulation studies have the potential to conclusively test if the motor system plays a causal role in beat perception and ground theories to their neural underpinnings.

Introduction

A tight relationship between movement and auditory rhythm perception is evident in human motor system response and motor involvement during music listening and rhythm tasks (Iversen & Balasubramaniam, 2016; Janata, Tomic, & Haberman, 2012; Repp, 2005a; Repp, 2005b; Ross, Warlaumont, Abney, Rigoli, & Balasubramaniam, 2016), and can be observed in neural response to music early in infant development (Kuhl, Ramirez, Bosseler, Lin, & Imada, 2014). How we move to music has by itself become a systematic subfield of inquiry (Ross et al., 2016) that often focuses on body synchronization with music. Some have suggested that the link between auditory and motor involvement in music could be similar to that found in language (Patel, Iversen, & Rosenberg, 2006).

Music often impels us to move in time with a perceived pulse or beat, implying a forward connection between auditory and motor systems that enables sound to guide movement planning and execution. Auditory training has been shown to improve motor performance (Stephan, Heckel, Song, & Cohen, 2015) and has even been explored for movement rehabilitation in patients with Parkinson's disease (Nombela, Hughes, Owen, & Grahn, 2013; Thaut et al., 1996) and recovery after stroke (Altenmuller & Schlaug, 2013). Interestingly, motor planning regions are active even when merely listening to music with a beat and not moving along. This raises the question: is the motor system necessary for beat perception, or is such motor activity a consequence of beat perception, reflecting unexecuted movement? The former view, perhaps surprising at first if one considers "the beat" to be a property of the music itself, is consonant with the idea that perception and movement are intimately coupled in a continuously interacting bidirectional perception–action relationship (Gibson, 1966). This perception–action relationship depends upon sensory perception to inform motor planning, but suggests that the motor system may influence active

perceptual processes, and this bidirectional causality is a characteristic of the models and theories reviewed here. While there is a long history of study in how the sensory systems inform action, there is now growing evidence that internal forward models make predictions about the sensory consequences of motor acts (Prinz, 1997; Wolpert & Flanagan, 2009). These predictions are thought to contribute to sensory perception and error assessments used for making corrections for discrepancies between expected and actual sensory input. An important task for music neuroscience is to understand to what extent there is a bidirectional relationship between auditory perception and action during not only performance but perception.

In this review, we focus on the particular relationship between motor planning and musical beat perception. We examine neural and behavioral evidence for active motor involvement in auditory rhythm perception, and contrast this with other general "motor" theories of action, and with motor theories of speech perception. In particular, we examine theories that posit the role of the motor system as "shadowing" or "mirroring" the auditory system and others that suggest that it may play a more causal or "predictive" role without which human musical beat perception would be impaired. In addition, we discuss theories that move beyond the more literal prediction of sensory consequences of motor acts to a more abstract role of the motor system in generating temporal predictions. We propose causal methods, such as using transcranial magnetic stimulation (TMS) protocols, as a necessary experimental step to further define the causal role of the motor system in auditory rhythm perception.

Beat perception

Beat perception refers to the detection of a regular pulse underlying a rhythmic input stream. Beats can be stressed, or accented, in regular patterns to help structure the pulse in a

predicable way, and this structure is referred to as meter (Lerdahl & Jackendoff, 1983; London, 2004). The auditory stream may be as simple as a metronome or as complex as a highly layered and time-varying musical work, but the human brain seems to almost automatically seek a simple regularity, the beat, or pulse, which can serve to organize our movements (as in dance, or tapping your foot to music), but also can organize our perception of time (Hannon, Snyder, Eerola, & Krumhansl, 2004; Palmer & Krumhansl, 1990). Two types of timing that are involved in rhythm perception are interval-based (absolute) timing and beat-based (relative) timing (Dalla Bella et al., 2016; Grube, Lee, Griffiths, Barker, & Woodruff, 2010; Iversen & Balasubramaniam, 2016). Interval-based timing is common to humans and non-human primates (Merchant & Honing, 2014; Zarco, Merchant, Prado, & Mendez, 2009). Beat-based timing may be uniquely human among primates (e.g., Merchant & Honing, 2014), and has other properties that make it of special interest for motor theories of timing (Grube et al., 2010; McAuley & Jones, 2003; Patel & Iversen, 2014).

A number of findings support the notion that motor activity may play an active role in shaping beat perception. In particular, as reviewed in the following, it has been suggested that beat-based timing relies on the establishment and maintenance of an internal predictive model, and there is support for this in demonstrations of anticipation in motor synchronization, tempo flexibility, the susceptibility of beat perception to willful control, and improved perceptual acuity of events that occur on the beat.

Many empirical studies have concluded that beat perception is anticipatory in nature (Miyake, 1902; Repp, 2005b; Woodrow, 1932). Analyses of finger tapping movements that are synchronized with an auditory rhythm demonstrate that taps often temporally precede the beat, an effect called negative mean asynchrony; humans spontaneously generate expectations of the timing of rhythmic components (Aschersleben, Gehrke, & Prinz, 2001; Drewing, Hennings, & Aschersleben, 2001). Another property of beat-based timing is flexibility in tempo. Rhythms can speed up or slow down and, despite these temporal fluctuations, people perceive an underlying rhythmic structure. In contrast to synchronous sound production in other species, which is not as demonstrably flexible, humans can entrain movements to a range of tempi between 94 and 176 BPM (Hanson, Case, Buck, & Buck, 1971; London, 2004; McAuley, Jones, Holub, Johnston, & Miller, 2006; Patel & Iversen, 2014; van Noorden & Moelants, 1999).

An important feature of musical beat is that it is a perceptual construct, influenced by but not uniquely determined by rhythms. It is susceptible to conscious control and active metrical interpretation on the part of the listener (Iversen, Repp, & Patel, 2009); the sense of beat actively shapes the perception of rhythm. In the aforementioned study (Iversen et al., 2009), when asked to impose different metrical interpretations onto a rhythmically ambiguous phrase, subjects' magnetoencephalography (MEG) recordings reflected imagined metrical structures despite physical stimulus invariance. Metrical interpretation influences early evoked neural responses in the beta range, with a stronger response on the imagined beat, and these patterns resemble those of non-imagined physical accents (Iversen et al., 2009).

Finally, beat perception is subject to influence by motor behavior. Overt body movement can improve perception of timing (Manning & Schutz, 2013) and influence perceptual interpretation of ambiguous rhythms (Phillips-Silver & Trainor, 2005, 2007). Overt and covert motor activities are associated with changes to perceptual acuity. Recent studies have shown improvements in beat-perception and finger-tapping entrainment to music when subjects were instructed to search for the pulse by moving their bodies (Su & Pöppel, 2012). Further research has shown that demonstrations of accelerating motion lead to faster perceived tempo of musical excerpts (Su, 2012; Su & Jonikaitis, 2011). Taken together, these results provide evidence that beat and meter perception are shaped by motor activity. However, in this review we will focus largely on the less explored role of the motor system in beat perception when no overt movement is involved.

Motor system activation during passive listening

A second strand of evidence suggesting the potential for a motor role in beat perception comes from neuroimaging, which has repeatedly shown that parts of the motor planning system are active during rhythm perception, even in the absence of overt movement, particularly for rhythms that evoke a strong sense of beat. In particular, beat perception engages dorsal premotor cortex, supplementary motor area (SMA), pre-SMA, basal ganglia, and lateral cerebellum (Bengtsson et al., 2009; Chen, Penhune, & Zatorre, 2008; Grahn & Brett, 2007; Stupacher, Hove, Novembre, Schütz-Bosbach, & Keller, 2013). Although these activations are distributed across multiple regions, they are all in areas of the brain associated with motor preparation and support that beat perception engages both motor and secondary motor structures.

What is motor activity doing when listening without moving? A parsimonious interpretation of this coactivation of motor regions while listening to rhythm is that it is related to anticipatory movement preparation. Consistent with this "mere" motor-preparation view, corticospinal excitability is modulated by listening to music with a strong beat. Stupacher et al. (2013) measured motor excitability during passive listening using TMS and demonstrated that excitability was time locked to the beat, and the degree of excitability reflected auditory–motor training, being greater in trained musicians. Motor excitability in amateur pianists while listening to a piano piece increases after learning to play that piece of music on the piano (D'Ausilio, Altenmüller, Belardinelli, & Lotze, 2006), and motor excitability is sensitive to differences in rhythmic properties between musical excerpts (Stupacher et al., 2013).

However, accumulating evidence suggests motor system activation while listening to rhythms may not merely be an epiphenomenon of suppressed movement, but may also play a causal role in shaping rhythm perception. One clue comes from cases where dysfunction of motor regions impairs rhythm perception. Patients with impaired basal ganglia function due to Parkinson's disease (PD) show impairments in a rhythm discrimination task compared with age-similar control subjects (Grahn & Brett, 2009). In the abovementioned study,

both PD and healthy participants were presented with a discrimination task using beat-based rhythms or non-beat-based rhythms. While there was no difference in discrimination ability between the groups for non-beat-based rhythms, discrimination of beat-based rhythms was reduced in the PD group suggesting that the basal ganglia are important for perception of musical beat, perhaps for generating an internal beat structure (Grahn & Rowe, 2013). Although further research is needed to determine how direct and causal the basal ganglia's involvement is in beat perception, recent work of Kotz, Brown, and Schwartze (2016) suggests that the basal ganglia's role in beat perception might stem from some aspect of motor preparation or planning.

Additional support for motor planning involvement in auditory perception can be found in MEG and electroencephalography studies of beta band neural oscillations. Beta band modulation is thought to be related to anticipatory processing, as beta activity decreases just after tone onset but its rebound may reflect sequence tempo (Fujioka, Trainor, Large, & Ross, 2012). The same study found coactivation of auditory and motor cortical areas, even without a motor response, although causality could not be addressed. Other results suggest that auditory beta-band modulation is influenced by top–down processes: voluntary metrical interpretation of rhythms modulates beta-band responses to sound (Iversen et al., 2009). Such modulation is suggested to reflect ongoing motor planning processes.

Motoric basis of sensory prediction

A third field of discourse that relates to the role of the motor system in beat perception is extensive work on internal motor models. From the perspective of this work, putative internal models used for auditory expectation could be understood as either *forward* or *inverse. Forward internal models*, such as efference copies, are used to predict sensory outcomes resulting from motor behavior. *Inverse internal models* are used to plan motor behavior based on desired sensory outcomes (Miall, 2003; Pfordresher, 2011; Tian & Poeppel, 2010). We can illustrate these types of models and their roles in coordinating action and perception with studies of singing. Studies of internal models involved in singing have used the impact of altering perceptual feedback to perturb ongoing production, presumably through the mismatch with an existing forward internal model and a disrupted inverse internal model creation. This has been shown with internal models for vocal pitch (Pfordresher, 2011) and volume production, known as the Lombard effect (Lombard, 1911; Zollinger & Brumm, 2011), and fingertip force production (Therrien, Lyons, & Balasubramaniam, 2012). Disruption of the process involving *inverse internal models* is thought to be the reason some people consistently sing pitches that are too high or low when trying to match pitch (Pfordresher, 2011). This phenomenon is referred to as poor-pitch singing, and has been shown to not be attributable to perceptual deficits in pitch perception, motor deficits, or pitch memory deficits (arguably, as outlined in Hutchins & Peretz, 2012). The majority of the general population can carry a tune with pitch and timing proficiency, but a small percentage cannot reliably match pitch even if they show no impairment in pitch discrimination tasks (Dalla Bella, Giguère, & Peretz, 2007). Pfordresher's (2011) explanation for poor-pitch singing is a deficit in creating an *inverse internal model* from perception of a pitch that can then be used for pitch production. This vocal imitation weakness has also been demonstrated in intonational speech (Pfordresher & Mantell, 2009). Accurate predictions about internal and external sensory effects of action are needed for skilled movement (Wolpert & Flanagan, 2009), including those underlying the articulatory processes in singing. In addition, predictions cannot be static but instead allow for online updating; errors between predictions and sensory consequences are continuously translated into changes in the internal model (Wolpert & Flanagan, 2015, 2009; Wolpert & Kawato, 1998; Yang, Wolpert, & Lengyel, 2016). Mere shadowing does not support this informative and flexible interchange between action and perceptual consequences via error-based modification to the internal model. There are clear advantages to having a bidirectional predictive basis to connect the auditory and motor systems.

Other domain-general frameworks have been developed, including ideomotor theories (Shin, Proctor, & Capaldi, 2010) and common-coding. Common-coding approaches present perception and action as having common representation (Prinz, 1997), thus making claims about the predictive nature of the relationship between the auditory and motor system. According to the common-coding accounts, actions are coded as the perceived effects of those actions (action effects). Thus viewing another person moving activates these action-perception representations, allowing for perceptual prediction generation. The theory is supported by evidence showing shared neural substrate for perceived and actualized movements, and by interference when the two try to access this representation simultaneously (Prinz, 1997), reminiscent of Gibson's (1966) account of perception being in service of generating opportunities for action and vice versa.

In a recent article, Press and Cook (2015) argued that domain-general motor contributions to perception undermine the theory that motor activation while watching human movement is for action simulation: that it is only shadowing. They describe a number of domain-general motor contributions to perception, including recognition of simple movement parameters (direction, position, velocity), timing, inference about human motion in masked point-light displays, mental rotation, and visual search. The authors classify these as domain-general because they are not necessarily related to complex motor actions such as grasping; these contributions could be understood as generically relating to movement. However, this evidence does not negate that these contributions are used to form sensory predictions.

Although there is considerable evidence for (Rizzolatti & Craighero, 2004) and controversies surrounding mirror neuron theories (Hickok, 2009), it has been argued that mirror neurons might play an important role in generating *inverse* and *forward internal models* (Miall, 2003). Mirror neurons are best known for their activity during visual observation, but can also become active when hearing an action without seeing the action (Kohler et al., 2002). Although there are obvious parallels between the theories of motor simulation and mirror

neurons (Koelsch, 2012), the mirror neurons associated with the ventral premotor area do not appear to be related directly to more dorsal premotor areas that are associated with beat perception.

Contrary to the mere shadowing accounts, the idea that the motor system may influence auditory cognition has been present in the literature for some time (Bolton, 1894), and has recently been discussed by a number of authors (Arnal, 2012; Jeannerod, 2001; Prinz, 1997; Rauschecker, 2011; Repp, 2005b; Schubotz, Friederici, & von Cramon, 2000; Sperry, 1952; Vuust, Ostergaard, Pallesen, Bailey, & Roepstorff, 2009; Zatorre, Chen, & Penhune, 2007). Recent accounts of sensory gain during movement support top–down motor influences on sensory state (Niell & Stryker, 2010; Nozaradan, Schönwiesner, Caron-Desrochers, & Lehmann, in press; Wekselblatt & Niell, 2015).

Andy Clark, Karl Friston and colleagues propose a predictive coding model that describes motor behavior as a way of selecting sensory input (Clark, 2015). In this framework, information flow is driven by top–down sensory predictions about proprioception and other sensory effects, and the only bottom up information is in the form of prediction errors. This active inference, or action-oriented predictive processing, proposes that downward connections from motor cortex (as from sensory cortex) carry predictions of sensory effects that are only met with bottom up prediction errors. This framework puts forth a forward model that is corrected when confronted by unexpected sensory consequences. It bypasses the need for inverse models and efference copies in favor of error modulated corollary discharge (encoded sensory predictions), and is a low cost strategy with minimal computational demands (Clark, 2015).

Another theory, the "ASAP" (Action Simulation for Auditory Prediction) hypothesis of Patel and Iversen (2014) makes a strong claim for a necessary predictive role of the motor system: activity in the motor planning system is necessary for beat-based perception, and fundamentally shapes our perception of events via connections in the dorsal auditory pathway enabling premotor, parietal and temporal cortices to interact. The ASAP hypothesis suggests that the motor planning system uses the same neural machinery involved in simulation of body movement (e.g., periodic movement patterns) to generate or entrain its neural activity patterns to the beat period, and that these patterns are communicated from motor planning regions to auditory regions where they serve as a predictive signal for the timing of upcoming beats and shape the perceptual interpretation of rhythms. This hypothesis expands on an earlier suggestion by Iversen et al. (2009) that in beat perception the motor system affects the auditory system by injecting precisely-timed beat related modulations, which itself was based on earlier psychological suggestion that the beat may involve "covert action" (Repp, 2005b). In contrast to "mirroring" theories, and to the motor theory of speech perception (discussed below), under ASAP the putative motor planning timing signals may, but need not, be related to imagery of movements of the type that would be required to create the perceptual input. Instead, they may be purely abstract timing, possibly, but not necessarily coupled to a specific action.

The central neuroscientific claim of the ASAP hypothesis is that beat perception involves temporally precise two-way communication between auditory regions and motor planning regions. This is related to the concept of reentry, "a process of temporally ongoing parallel signaling between separate maps along ordered anatomical connections" (Edelman, 1989). According to ASAP, (1) neural signals from auditory to motor planning regions provide information about the timing of auditory events; (2) these signals influence the timing of periodic motor planning signals in motor regions, and (3) these planning signals flow from motor regions back to auditory regions to provide a signal that predicts upcoming beat times. In forward models such as predictive coding, primary information processing operates on predictions of sensory consequences, but in simulation-based models such as ASAP, top–down (anticipatory) and bottom–up (reactive) processes work in parallel, continuously influencing each other.

Other extant models of beat perception that posit top–down influences on auditory processing include dynamic attending theory where attention is modulated with temporal event structure (Jones, 1976), and hypotheses relating motor influence on auditory processing to active suppression during vocalization (Arnal, 2012). Nonlinear oscillator models suggest one way this might be achieved by entrainment of hypothesized neural oscillations with rhythmic auditory events with reciprocal interactions among several layers of the network required to predict the beat (Large, Herrera, & Velasco, 2015; Large & Jones, 1999; Large & Snyder, 2009).

The path ahead

The evidence reviewed above can be organized into two perspectives: motor system activation while listening to rhythms is (1) only shadowing or (2) it also has a predictive, causal role in beat perception. Much of the evidence is suggestive of a causal role, but many questions remain that need to be answered to move forward with this work. How might we make further progress on these questions? Below we remark on other motor theories of perception and how they do and do not contribute to advancing support of either of these perspectives. We then make experimental suggestions for causal studies needed to test the validity of these perspectives. Further work is required to clearly elucidate the role of the motor planning regions in the auditory dorsal stream that could also help distinguish between how the brain responds differentially to music and other acoustic stimulation like speech (Hickok, Buchsbaum, Humphries, & Muftuler, 2003).

Motor simulation theories

There is a long history of discussing motor involvement in speech perception, and we might turn to it for comparison and contrast. Much as there is motor activation when listening to rhythms, numerous studies show neuroimaging evidence of motor activation while participants listen to speech (Skipper, Nusbaum, & Small, 2005; Wilson, Saygin, Sereno, & Iacoboni, 2004). In addition, there is a range of evidence to suggest that speech effector muscles show facilitation when listening to speech (Fadiga, Craighero, Buccino, & Rizzolatti, 2002). MEG analysis of infants at ages 7–11 months supports that motor

activation while listening to speech is present in infants that are just learning to make pre-speech sounds (Kuhl et al., 2014). It would appear that this multisensory and multimodal relationship is present very early in speech development, and that it is likely an integral part of speech perception.

Some accounts of speech perception, such as the Motor Theory of Speech Perception (Galantucci, Fowler, & Turvey, 2006; Liberman & Mattingly, 1985) and Analysis by Synthesis (Stevens & Halle, 1967), propose that speech perception *relies on* motor estimation or expectation. The Motor Theory of Speech Perception was proposed to address the problem of perceptual invariance (Liberman & Mattingly, 1985), which makes its motivation quite distinct from motor simulation theories of rhythm perception. Invariance is the observation that a speech signal can have considerable acoustic variation due to context, the speaker's gender and vocal qualities, and noise but listeners are still able to group speech sounds into meaningful categories such as phonemes. Liberman and colleagues proposed that speech sound categories are derived by inferring the neural representation of the gestures that produced the sound (Liberman & Mattingly, 1985).

The Motor Theory of Speech Perception has been criticized for a number of reasons including its relation to theories of speech modularity and lack of specification about how acoustic signals are mapped to gestures (Sussman, 1989). Another criticism of the Motor Theory of Speech Perception and motor theories in general is that patients with damage to the motor system can exhibit normal action recognition (Stasenko, Garcea, & Mahon, 2013), and normal phonemic discrimination, although explicit labeling of speech sounds is impaired (Stasenko et al., 2015). Stasenko et al., posit a more nuanced view that motor representations may be called upon when other language cues are not present (such as semantics or context).

Although there are similarities between motor theories for speech perception and motor theories for beat perception, there are also many clear differences. Perhaps foremost, in motor theories of speech, sounds are mapped to motor representations the perceiver would use to produce the same sounds. In complex music, at least for non-musicians, such a direct mirroring is inconceivable. Speech perception relies on linguistic context in a way that beat perception does not. Second, as mentioned above, the motivations behind the theories are distinct. Motor activity in speech perception was proposed as a mechanism for creating speaker invariance of speech perception, whereas motor involvement in rhythm perception is proposed as a source of temporally precise signals to modulate rhythmic expectation and grouping, as well as implement the observed willful endogenous influences on rhythm perception. Third, the neural circuits implicated in motor theories of speech are distinct from those proposed by motor simulation theories of beat perception. Speech listening has been shown to be accompanied by bilateral activations in superior ventral premotor cortex, which are associated with speech motor production, and in primary motor cortex (Wilson et al., 2004). Although there is activation during passive speech listening in motor areas (Wilson, Molnar-Szakacs, & Iacoboni, 2008; Wilson et al., 2004), this does not provide support for bidirectional predictive auditory–motor processes in the dorsal auditory stream. In addition, the temporal regularity of rhythmic contexts could enable prediction in a way that naturalistic speech might not. Beats have a more predictable structure in a way that speech, with all its irregularities, does not. For this reason, beat perception might allow for more motor simulation than speech perception, and beat perception paradigms might be more ideal for investigations of perceptually relevant motor recruitment.

Causal studies of neural circuits during beat perception

TMS is a technique that uses magnetic field pulses applied to the surface of the scalp to cause functional changes in the electrical neural activity in superficial cortical regions (Huang, Edwards, Rounis, Bhatia, & Rothwell, 2005). TMS protocols use parameters such as magnetic field power and pulse frequency to induce temporary excitement or disruption to target cortical regions, and therefore can be used to change cortical activity during or before asking participants to do behavioral experiments to see if changing cortical activation leads to changes in behavior (Huang et al., 2005).

TMS protocols have been used to explore motor theories of speech. TMS-induced disruption of premotor cortex has been shown to disrupt speech perception (Meister, Wilson, Deblieck, Wu, & Iacoboni, 2007), but Stasenko and colleagues argue that the spread of TMS-induced changes might lead to disruption of sensory regions in addition to premotor targets. Figure eight TMS coil designs advertise focal stimulation, but the spread of activation is not well understood. However, even in studies of clinical lesions we see support of motor theories.

There is some existing work with TMS showing, using causal designs, the neural substrate involved with timing abilities. Low-frequency repetitive TMS applied over left dorsolateral premotor cortex (dPMC) can interfere with accuracy on a finger tapping synchronization task to an auditory metronome, and this disruption in accuracy occurs whether the participant is tapping with their right or left hand (Pollok, Rothkegel, Schnitzler, Paulus, & Lang, 2008). Because synchronization employs beat-based predictive timing mechanisms, it may be concluded that left dPMC is involved in beat-based timing. Continuous theta burst stimulation (cTBS), a TMS protocol that down-regulates cortical activity at the focal target location (Huang et al., 2005), interferes with interval-based timing when applied over medial cerebellum, but does not interfere with beat-based timing (Grube et al., 2010). This supports a functional dissociation between interval and beat-based timing, and suggests that cerebellum is involved in interval, but not beat-based, timing.

Although there is a scarcity of causal work on neural contributors to beat-based timing, the weight of these studies is considerable due to strengths of the causal designs. Additional causal work is needed to explore the current motor theories of beat perception, and to ground these theories to their neural underpinnings. TMS protocols provide powerful causal methodology that can temporarily alter cortical activity in focal motor and premotor regions, either by exciting or suppressing activity (Huang et al., 2005), and this can be used to test theories that claim that the role of the motor system is obligatory for beat-

based timing. For example, disruption of the internal model/simulation mechanisms for beat perception should lead to declines in accurate rhythm perception and auditory–motor synchronization. Thus, manipulating motor planning activity and internal model generation in a rhythm task should lead to changes in accurate rhythm perception and production.

Conclusions

The theories related to speech perception suggest that motor activation during speech listening could be mirroring/shadowing used for auditory processing. The long history of motor theories of speech provide beat perception theories with frameworks in which to operate, but do not meet the standard of proof for theories that propose that motor activation during music listening reflects predictive processes. Beat-perception and entrainment paradigms, internal model frameworks, and demonstrations of functional connectivity are contributing to mounting evidence for predictive simulations in motor networks. However, these methodologies can only provide suggestive evidence. Extending these paradigms using causal methodology is needed in order to conclusively show that motor networks not only shadow speech and music but also provide predictive models that can be actively updated and maintained. Direct tests of a causal role of the motor system in beat perception are needed that perturb the motor system, either through dual task paradigms or through direct neurostimulation. Motor theories of perception, such as predictive coding and the ASAP hypothesis, propose that motor simulation is likely integral to auditory beat-based timing. Beat perception studies have been used to demonstrate the strong relationship between motor activation and listening to repetitive sounds, and this relationship has possible implications for understanding evolutionary origins of music and its relation to language. Tests of the theories of motor involvement for speech perception provide support for robustness of the link between motor system activation and auditory processing, and with evidence from beat-perception and internal model paradigms, provide suggestive evidence for predictive simulation. Further research should implement causal studies to directly test for predictive motor models in beat perception. TMS is a technique that can be used non-invasively to investigate the causal relationship between motor simulation and beat-based timing, and more generally help explore beat perception as a specific example of an action-based perceptual system, providing an example of how the study of music and brain can address general mechanisms of brain function.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

ORCID

Ramesh Balasubramaniam http://orcid.org/0000-0001-7575-3080

References

Altenmuller, E., & Schlaug, G. (2013). Neurologic music therapy: The beneficial effects of music making on neurorehabilitation. *Acoustical Science and Technology*, *34*, 5–12. doi:10.1250/ast.34.5

Arnal, L. H. (2012). Predicting "when" using the motor system's beta-band oscillations. *Frontiers in Human Neuroscience*, *6*, 225. doi:10.3389/fnhum.2012.00225

Aschersleben, G., Gehrke, J., & Prinz, W. (2001). Tapping with peripheral nerve block. A role for tactile feedback in the timing of movements. *Experimental Brain Research*, *136*, 331–339. doi:10.1007/s002210000562

Bengtsson, S. L., Ullén, F., Ehrsson, H. H., Hashimoto, T., Kito, T., Naito, E. … Sadato, N. (2009). Listening to rhythms activates motor and premotor cortices. *Cortex*, *45*, 62–71. doi:10.1016/j.cortex.2008.07.002

Bolton, T. (1894). Rhythm. *The American Journal of Psychology*, *6*, 145–238. doi:10.2307/1410948

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex*, *18*, 2844–2854. doi:10.1093/cercor/bhn042

Clark, A. (2015). Embodied prediction. In T. Metzinger & J. M. Windt (Eds.), *Open MIND* (Vol. 7). Frankfurt am Main: MIND Group. doi:10.15502/9783958570115

D'Ausilio, A., Altenmüller, E., Belardinelli, M. O., & Lotze, M. (2006). Cross-modal plasticity of the motor cortex while listening to a rehearsed musical piece. *European Journal of Neuroscience*, *24*, 995–958. doi:10.1111/j.1460-9568.2006.04960.x

Dalla Bella, S., Farrugia, N., Benoit, C.-E., Begel, V., Verga, L., Harding, E., & Kotz, S. A. (2016). BAASTA: Battery for the assessment of auditory sensorimotor and timing abilities. *Behavior Research Methods*. doi:10.3758/s13428-016-0773-6

Dalla Bella, S., Giguère, J.-F., & Peretz, I. (2007). Singing proficiency in the general population. *The Journal of the Acoustical Society of America*, *121*, 1182–1189. doi:10.1121/1.2427111

Drewing, K., Hennings, M., & Aschersleben, G. (2001). The contribution of tactile reafference to temporal regularity during bimanual finger tapping. *Psychological Research*, *66*, 60–70. doi:10.1007/s004260100074

Edelman, C. M. (1989). *The remembered present: A biological theory of consciousness*. New York, NY: Basic Books.

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience*, *15*, 399–402. doi:10.1046/j.0953-816x.2001.01874.x

Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2012). Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. *Journal of Neuroscience*, *32*, 1791–1802. doi:10.1523/JNEUROSCI.4107-11.2012

Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, *13*, 361–377. doi:10.3758/BF03193857

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston, MA: Houghton Mifflin.

Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, *19*, 893–906. doi:10.1162/jocn.2007.19.5.893

Grahn, J. A., & Brett, M. (2009). Impairment of beat-based rhythm discrimination in Parkinson's disease. *Cortex*, *45*, 54–61. doi:10.1016/j.cortex.2008.01.005

Grahn, J. A., & Rowe, J. B. (2013). Finding and feeling the musical beat: Striatal dissociations between detection and prediction of regularity. *Cerebral Cortex*, *23*, 913–921. doi:10.1093/cercor/bhs083

Grube, M., Lee, K.-H., Griffiths, T. D., Barker, A. T., & Woodruff, P. W. (2010). Transcranial magnetic theta-burst stimulation of the human cerebellum distinguishes absolute, duration-based from relative, beat-based perception of subsecond time intervals. *Frontiers in Psychology*, *1*, 171. doi:10.3389/fpsyg.2010.00171

Hannon, E. E., Snyder, J. S., Eerola, T., & Krumhansl, C. L. (2004). The role of melodic and temporal cues in perceiving musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 956–974. doi:10.1037/0096-1523.30.5.956

Hanson, F. E., Case, J. F., Buck, E., & Buck, J. (1971). Synchrony and flash entrainment in a New Guinea firefly. *Science*, 174, 162–164. doi:10.1126/science.174.4005.161

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*, 21, 1229–1243. doi:10.1162/jocn.2009.21189

Hickok, G., Buchsbaum, B., Humphries, C., & Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience*, 15, 673–682. doi:10.1162/089892903322307393

Huang, Y., Edwards, M. J., Rounis, E., Bhatia, K. P., & Rothwell, J. C. (2005). Theta burst stimulation of the human motor cortex. *Neuron*, 45, 201–206. doi:10.1016/j.neuron.2004.12.033

Hutchins, S., & Peretz, I. (2012). A frog in your throat or in your ear? Searching for the causes of poor singing. *Journal of Experimental Psychology: General*, 141, 76–97. doi:10.1037/a0025064

Iversen, J. R., & Balasubramaniam, R. (2016). Synchronization and temporal processing. *Current Opinion in Behavioral Sciences*, 8, 175–180. doi:10.1016/j.cobeha.2016.02.027

Iversen, J. R., Repp, B. H., & Patel, A. D. (2009). Top-down control of rhythm perception modulates early auditory responses. *Annals of the New York Academy of Sciences*, 1169, 58–73. doi:10.1111/j.1749-6632.2009.04579.x

Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology: General*, 141, 54–75. doi:10.1037/a0024208

Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism for motor cognition. *NeuroImage*, 14, S103–109. doi:10.1006/nimg.2001.0832

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 83, 323–355. doi:10.1037/0033-295X.83.5.323

Koelsch, S. (2012). *Brain and music*. West Sussex: Wiley-Blackwell.

Kohler, E., Keysers, C., Umilt`a, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: Action representation in mirror neurons. *Science*, 297, 846–848. doi:10.1126/science.1070311

Kotz, S. A., Brown, R. M., & Schwartze, M. (2016). Cortico-striatal circuits and the timing of action and perception. *Current Opinion in Behavioral Sciences*, 8, 42–45. doi:10.1016/j.cobeha.2016.01.010

Kuhl, P. K., Ramirez, R. R., Bosseler, A., Lin, J. L., & Imada, T. (2014). Infants' brain responses to speech suggest analysis by synthesis. *Proceedings of the National Academy of Sciences*, 111, 11238–11245. doi:10.1073/pnas.1410963111

Large, E., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, 106, 119–159. doi:10.1037/0033-295X.106.1.119

Large, E. W., Herrera, J. A., & Velasco, M. J. (2015). Neural networks for beat perception in musical rhythm. *Frontiers in Systems Neuroscience*, 9, 583. doi:10.3389/fnsys.2015.00159

Large, E. W., & Snyder, J. S. (2009). Pulse and meter as neural resonance. *Annals of the New York Academy of Sciences*, 1169, 46–57. doi:10.1111/j.1749-6632.2009.04550.x

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge: MIT Press.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21, 1–36. doi:10.1016/0010-0277(85)90021-6

Lombard, E. (1911). Le signe de l'élévation de la voix. *Annales des Maladies de L'Oreille et du Larynx*, 37, 101–119.

London, J. (2004). *Hearing in time*. New York, NY: Oxford University Press.

Manning, F., & Schutz, M. (2013). "Moving to the beat" improves timing perception. *Psychonomic Bulletin & Review*, 20, 1133–1139. doi:10.3758/s13423-013-0439-7

McAuley, J. D., & Jones, M. R. (2003). Modeling effects of rhythmic context on perceived duration: a comparison of interval and entrainment approaches to short-interval timing. *Journal of Experimental Psychology: Human Perception and Performance*, 29, 1102–1125. doi:10.1037/0096-1523.29.6.1102

McAuley, J. D., Jones, M. R., Holub, S., Johnston, H. M., & Miller, N. S. (2006). The time of our lives: Life span development of timing and event tracking. *Journal of Experimental Psychology General*, 135, 348–367. doi:10.1037/0096-3445.135.3.348

Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., & Iacoboni, M. (2007). The essential role of premotor cortex in speech perception. *Current Biology*, 17, 1692–1696. doi:10.1016/j.cub.2007.08.064

Merchant, H., & Honing, H. (2014). Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Frontiers in Psychology*, 7, 274. doi:10.3389/fnins.2013.00274

Miall, R. C. (2003). Connecting mirror neurons and forward models. *NeuroReport*, 14, 2135–2137. doi:10.1097/00001756-200312020-00001

Miyake, I. (1902). Researches on rhythmic activity. *Studies From the Yale Psychological Laboratory*, 10, 1–48.

Niell, C. M., & Stryker, M. P. (2010). Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron*, 65, 472–479. doi:10.1016/j.neuron.2010.01.033

Nombela, C., Hughes, L. E., Owen, A. M., & Grahn, J. A. (2013). Into the groove: Can rhythm influence Parkinson's disease? *Neuroscience & Biobehavioral Reviews*, 37, 2564–2570. doi:10.1016/j.neubiorev.2013.08.003

Nozaradan, S., Schönwiesner, M., Caron-Desrochers, L., & Lehmann, A. (in press). Enhanced brainstem and cortical encoding of sound during synchronized movement. *NeuroImage*. doi:10.1016/j.neuroimage.2016.07.015

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 728–741. doi. 10.1037/0096-1523.16.4.728

Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: The Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Human Neuroscience*, 8, 57. doi:10.3389/fnsys.2014.00057

Patel, A. D., Iversen, J. R., & Rosenberg, J. C. (2006). Comparing the rhythm and melody of speech and music: The case of British English and French. *The Journal of the Acoustical Society of America*, 119, 3034–3047. doi:10.1121/1.2179657

Pfordresher, P. (2011). *Poor-pitch singing as an inverse model deficit: Imitation and estimation*. International Symposium on Performance Science, pp. 539–544.

Pfordresher, P. Q., & Mantell, J. (2009). Singing as a form of vocal imitation: Mechanisms and deficits. *Proceedings of the 7th triennial conference of European Society for the Cognitive Sciences of Music*, 425–430.

Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science*, 308, 1430. doi:10.1126/science.1110922

Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition*, 105, 533–546. doi:10.1016/j.cognition.2006.11.006

Pollok, B., Rothkegel, H., Schnitzler, A., Paulus, W., & Lang, N. (2008). The effect of rTMS over left and right dorsolateral premotor cortex on movement timing of either hand. *European Journal of Neuroscience*, 27, 757–764. doi:10.1111/j.1460-9568.2008.06044.x

Press, C., & Cook, R. (2015). Beyond action-specific simulation: Domain-general motor contributions to perception. *Trends in Cognitive Sciences*, 19, 176–178. doi:10.1016/j.tics.2015.01.006

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, 9, 129–154. doi:10.1080/713752551

Rauschecker, J. P. (2011). An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hearing Research*, 271, 16–25. doi:10.1016/j.heares.2010.09.001

Repp, B. H. (2005a). Rate limits of on-beat and off-beat tapping with simple auditory rhythms: 2. The roles of different kinds of accent. *Music Perception*, 23, 165–187. doi:10.1525/mp.2005.23.2.165

Repp, B. H. (2005b). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, 12, 969–992. doi:10.3758/BF03206433

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169–192. doi:10.1146/annurev.neuro.27.070203.144230

Ross, J. M., Warlaumont, A. S., Abney, D. H., Rigoli, L. M., & Balasubramaniam, R. (2016). Influence of musical groove on postural sway. *Journal of Experimental Psychology: Human Perception and Performance*. Advance online publication. doi:10.1037/xhp0000198

Schubotz, R. I., Friederici, A. D., & von Cramon, D. Y. (2000). Time perception and motor timing: A common cortical and subcortical basis revealed by fMRI. *NeuroImage*, 11, 1–12. doi:10.1006/nimg.1999.0514

Shin, Y. K., Proctor, R. W., & Capaldi, E. J. (2010). A review of contemporary ideomotor theory. *Psychological Bulletin*, 136, 943–974. doi:10.1037/a0020541

Skipper, J. I., Nusbaum, H. C., & Small, S. L. (2005). Listening to talking faces: Motor cortical activation during speech perception. *NeuroImage*, 25, 76–89. doi:10.1016/j.neuroimage.2004.11.006

Sperry, R. W. (1952). Neurology and the mind-brain problem. *American Scientist*, 40(2).

Stasenko, A., Bonn, C., Teghipco, A., Garcea, F. E., Sweet, C., Dombovy, M. … Mahon, B. Z. (2015). A causal test of the motor theory of speech perception: A case of impaired speech production and spared speech perception. *Cognitive Neuropsychology*, 32, 38–57. doi:10.1080/02643294.2015.1035702

Stasenko, A., Garcea, F. E., & Mahon, B. Z. (2013). What happens to the motor theory of perception when the motor system is damaged? *Language and Cognition*, 5, 225–238. doi:10.1515/langcog-2013-0016

Stephan, M. A., Heckel, B., Song, S., & Cohen, L. G. (2015). Crossmodal encoding of motor sequence memories. *Psychological Research*, 79, 318–326. doi:10.1007/s00426-014-0568-2

Stevens, K. N., & Halle, M. (1967). Remarks on analysis by synthesis and distinctive features. In W. Wathem-Dunn (Ed.), *Models for the perception of speech and visual form*. Cambridge, MA: MIT Press.

Stupacher, J., Hove, M. J., Novembre, G., Schütz-Bosbach, S., & Keller, P. E. (2013). Musical groove modulates motor cortex excitability: A TMS investigation. *Brain and Cognition*, 82, 127–136. doi:10.1016/j.bandc.2013.03.003

Su, Y.-H. (2012). *The influence of external and internal motor processes on human auditory rhythm perception*. (Doctoral dissertation). Aus dem Institut für Medizinische Psychologie der Ludwig-Maximilians-Universität, München.

Su, Y.-H., & Jonikaitis, D. (2011). Hearing the speed: Visual motion biases the perception of auditory tempo. *Experimental Brain Research*, 214, 357–371. doi:10.1007/s00221-011-2835-4

Su, Y.-H., & Pöppel, E. (2012). Body movement enhances the extraction of temporal structures in auditory sequences. *Psychological Research*, 76, 373–382. doi:10.1007/s00426-011-0346-3

Sussman, H. M. (1989). Neural coding of relational invariance in speech: Human language analogs to the barn owl. *Psychological Review*, 96, 631–642. doi:10.1037/0033-295X.96.4.631

Thaut, M. H., McIntosh, G. C., Rice, R. R., Miller, R. A., Rathbun, J., & Brault, J. M. (1996). Rhythmic auditory stimulation in gait training for Parkinson's disease patients. *Movement Disorders*, 11, 193–200. doi:10.1002/mds.870110213

Therrien, A. S., Lyons, J., & Balasubramaniam, R. (2012). Sensory attenuation Q15 of self-produced feedback: The Lombard effect revisited. *PLoS One*, 7, e49370. doi:10.1371/journal.pone.0049370

Tian, X., & Poeppel, D. (2010). Mental imagery of speech and movement implicates the dynamics of internal forward models. *Frontiers in Psychology*, 1–166. doi:10.3389/fpsyg.2010.00166

van Noorden, L., & Moelants, D. (1999). Resonance in the perception of musical pulse. *Journal of New Music Research*, 28, 43–66. doi:10.1076/jnmr.28.1.43.3122

Vuust, P., Ostergaard, L., Pallesen, K. J., Bailey, C., & Roepstorff, A. (2009). Predictive coding of music: Brain responses to rhythmic incongruity. *Cortex*, 45, 80–92. doi:10.1016/j.cortex.2008.05.014

Wekselblatt, J. B., & Niell, C. M. (2015). Behavioral state—Getting "In The Zone.". *Neuron*, 87, 7–9. doi:10.1016/j.neuron.2015.06.020

Wilson, S. M., Molnar-Szakacs, I., & Iacoboni, M. (2008). Beyond superior temporal cortex: Intersubject correlations in narrative speech comprehension. *Cerebral Cortex*, 18, 230–242. doi:10.1093/cercor/bhm049

Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, 7, 701–702. doi:10.1038/nn1263

Wolpert, D. M., & Flanagan, J. R. (2009). Forward models. In T. Baynes, A. Cleeremans, & P. Wilken (Eds.), *The oxford companion to consciousness*. Oxford: Oxford University Press.

Wolpert, D. M., & Flanagan, J. R. (2015). Computations underlying sensorimotor learning. *Current Opinion in Neurobiology*, 37, 7–11. doi:10.1016/j.conb.2015.12.003

Wolpert, D. M., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, 11, 1317–1329. doi:10.1016/S0893-6080(98)00066-5

Woodrow, H. (1932). The effect of rate of sequence upon the accuracy of synchronization. *Journal of Experimental Psychology*, 15, 357–379. doi:10.1037/h0071256

Yang, S. C., Wolpert, D. M., & Lengyel, M. (2016). Theoretical perspectives on active sensing. *Current Opinion in Behavioral Sciences*, 11, 100–108. doi:10.1016/j.cobeha.2016.06.009

Zarco, W., Merchant, H., Prado, L., & Mendez, J. C. (2009). Subsecond timing in primates: Comparison of interval production between human subjects and rhesus monkeys. *Journal of Neurophysiology*, 102, 3191–3202. doi:10.1152/jn.00066.2009

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience*, 8, 547–558. doi:10.1038/nrn2152

Zollinger, S. A., & Brumm, H. (2011). The Lombard effect. *Current Biology*, 21, R614–615. doi:10.1016/j.cub.2011.06.003