

Vestibular influence on auditory metrical interpretation

Jessica Phillips-Silver, Laurel J. Trainor *

Department of Psychology, Neuroscience and Behaviour, McMaster University, Hamilton, Ont., Canada L8S 4K1

Accepted 29 November 2007
Available online 29 January 2008

Abstract

When we move to music we feel the beat, and this feeling can shape the sound we hear. Previous studies have shown that when people listen to a metrically ambiguous rhythm pattern, moving the body on a certain beat—adults, by actively bouncing themselves in synchrony with the experimenter, and babies, by being bounced passively in the experimenter's arms—can bias their auditory metrical representation so that they interpret the pattern in a corresponding metrical form [Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science*, 308, 1430; Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition*, 105, 533–546]. The present studies show that in adults, as well as in infants, metrical encoding of rhythm can be biased by passive motion. Furthermore, because movement of the head alone affected auditory encoding whereas movement of the legs alone did not, we propose that vestibular input may play a key role in the effect of movement on auditory rhythm processing. We discuss possible cortical and subcortical sites for the integration of auditory and vestibular inputs that may underlie the interaction between movement and auditory metrical rhythm perception.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Auditory system; Rhythm; Music; Vestibular system; Metrical structure; Movement

1. Introduction

Music engenders movement. The idea that music and physical motion are related is found across cultures and can be traced to antiquity (e.g., Todd, 1995; Todd, Cousins, & Lee, 2007). Musical ideas are often expressed by using movement as a metaphor, such as in expressions of speed and timing: “the music is *slowing down*”, or “the tempo is *andante con moto*” (*walking speed with movement*), expressions of shape and form: “*moving from note to note*” or “*tracing an arabesque*”, and expressions of style: “*a flowing melody*”, or “*a swinging rhythm*” (e.g., Driver, 1936; Shove & Repp, 1995). While such metaphors have artistic value, a number of researchers have also proposed that the tie between music and movement is concrete (e.g., Clarke, 1993; Clarke, 1999; Fraisse, 1982; Palmer, 1997; Todd, 1999), and that, indeed, music may have evolved from physical movement (e.g., Todd et al., 2007; Trainor,

2007). Evidence for this idea comes from biomechanical studies of the similarity between locomotion rate and preferred musical tempo (e.g., Todd et al., 2007); from studies showing that people readily produce movements in time to music (Repp, 2005; Repp & Doggett, 2007); from studies showing that people judge with ease the dance movements of others (Brown et al., 2005); and from imaging studies showing that brain regions responsible for synchronized movement are modulated by auditory metrical structure (e.g., Chen, Zatorre, & Penhune, 2006; Haueisen & Knosche, 2001; Lahav, Saltzman, & Schlaug, 2007; Zatorre, Chen, & Penhune, 2007). The relation between movement and musical rhythm may also have important developmental functions (Phillips-Silver & Trainor, 2005), and has been proposed to drive interpersonal synchrony between mothers and infants (Trehub, 2003).

Interestingly, not only does musical rhythm activate motor areas in the brain and make us want to move, but movement can enhance listening. Musicians often use expressive body gestures to convey the timing or the feeling of motion in the music (Thompson, Graham, & Russo,

* Corresponding author. Fax: +1 905 529 6225.
E-mail address: LJT@mcmaster.ca (L.J. Trainor).

2005). Basic and advanced teaching of rhythm in music typically involves first feeling the beat in the body through overt body movements, and then internalizing the rhythm in an auditory code that may well involve a covert motor representation (Jaques-Dalcroze, 1920; Juntunen & Hyvönen, 2004). The experienced listener may no longer need to actually move in order to hear the beat, but rhythm perception may nonetheless originate in movement. Our previous behavioral studies support the idea that movement can affect auditory processing of rhythm, showing that body movement can disambiguate a metrically ambiguous rhythmic sound pattern (Phillips-Silver & Trainor, 2005; Phillips-Silver & Trainor, 2007). We trained infant and adult participants to listen to a rhythmic drumbeat pattern with no accented strong beats (i.e., no imposed metrical structure), repeating for a period of two minutes. While they heard the rhythm pattern, we had them bounce to the rhythm in one of two ways: either on every second beat, or on every third beat. Participants were then tested with two auditory-only versions of the rhythm pattern: one with every second beat accented (i.e., duple version), and the other with every third beat accented (i.e., triple version). We recorded infants' listening times to each test stimulus in a preference procedure. Adults were asked to identify which test stimulus matched what they had heard during training. Participants of both age groups identified the test stimulus that matched the metrical form in which they had bounced: participants who had bounced in duple form recognized as familiar the duple version of the rhythm pattern, while those who had bounced in triple form recognized as familiar the triple version of the rhythm pattern. We also observed the movement effect with participants who were blindfolded during training, indicating that visual input is not necessary for the effect. However we observed no effect with participants who sat and passively observed as the experimenter bounced during training, indicating that movement of one's own body is critical. We concluded that the way we move can shape what we hear.

While our previous studies (Phillips-Silver & Trainor, 2005, 2007) demonstrate a multisensory interaction between movement and auditory perception, they do not indicate which aspects of movement are critical to the effect. Body movement involves several different kinds of sensorimotor information, including motor planning, proprioception, tactile, and vestibular inputs (Nolte, 2002). These inputs can vary with the type of movement experienced (e.g., full-body rotation versus head tilt; active versus passive movement) (Cullen & Minor, 2002; Cullen & Roy, 2004; Goldberg & Fernandez, 1980; Klam & Graf, 2003; Klinke & Schmidt, 1970). In particular, the vestibular system is responsible for detecting motion of the head in space, which contributes to spatial perception, allows for maintaining appropriate body orientation, and is crucial for functioning in the environment (Gdowski & McCrea, 1999). Because infants showed the multisensory interaction with passive motion (i.e., they were bounced in the arms of an adult), and because the vestibular system develops early

and influences motor development in infants (Clark, Kretzberg, & Chee, 1977; Romand, 1992; Shahidullah & Hepper, 1994), the present experiments investigated whether vestibular input is sufficient to induce the auditory–movement effect. Specifically, the present experiments were designed to reduce or eliminate motor, proprioceptive and tactile information, thereby isolating the effect of vestibular input on auditory metrical disambiguation.

Infants in the previous studies were moved passively in the arms of the experimenter whereas adults generated their own movement (Phillips-Silver & Trainor, 2005, 2007). In the present experiments we first asked whether passive motion in adults could bias their metrical interpretation of the auditory rhythm pattern. If passive motion is sufficient, the vestibular system is likely to play a critical role in the interaction between movement and audition because motor planning, proprioception, and tactile input are all reduced under passive compared to active movement. We then tested for vestibular involvement by observing whether head movement alone, which activates the vestibular system, and lower body movement alone, which does not activate the vestibular system, can bias the metrical interpretation of an auditory rhythm pattern.

To summarize the goals of the present study, we asked two questions: (1) whether adults' metrical encoding of the rhythm relies on active, self-generated movement, or whether passive motion is sufficient, and (2) whether head motion alone is sufficient to cause the multisensory interaction, thereby implicating a role of the vestibular system.

2. Experiment 1

In our previous studies, adults who moved actively by bending their knees and bouncing up and down, either on every second beat or on every third beat of an ambiguous six-beat rhythm pattern, were biased to encode the auditory stimulus in either duple or triple form, respectively. The goal of the present study was to observe whether passive motion also results in biasing adults' metrical interpretation of the ambiguous rhythm pattern.

2.1. Method

2.1.1. Participants

The study included 8 healthy university undergraduate students (aged 19–32 years, mean 22.4 years). Musical training (defined by past or present lessons in musical instruments, voice or dance) ranged from 0 to 15 years (mean 9.3 years). Subjects reported whether or not they participate in any recreational (i.e., without training) music or dancing, either private or public (such as dancing in night clubs): all 8 reported some recreational music activity; 4 of the 8 reported recreational dancing. In all experiments subjects had no known hearing deficits. Procedures were approved by the McMaster University Research Ethics Board, and adults in all studies gave written consent to participate.

2.1.2. Stimuli

The stimuli were identical to those of Phillips-Silver and Trainor (2005) (see <http://www.sciencemag.org/cgi/content/full/308/5727/1430> to hear experimental stimuli).

2.1.2.1. Training stimulus. The training stimulus was constructed as follows (see Fig. 1), and presented in a sound-attenuating chamber with a noise floor of 29 dB. First, a downbeat and microbeat were introduced. The *downbeat* (snare drum timbre) was presented at 60 dB with an SOA of 1998 ms. After four repetitions, a *microbeat* (slapstick timbre) background with an SOA of 333 ms, presented at 50 dB, began and repeated throughout the rest of the training stimulus presentation. This downbeat plus microbeat combination resulted in a six-beat background sequence, with the snare drum sounding on the first beat, followed by five slapstick beats. The presence, and relative loudness, of the snare drum downbeat helped to perceptually divide the beat pattern into measures, or musical bars of six beats each. This combination was presented for eight measures. Next, the *training rhythm* of interest was superimposed. The training rhythm was the same duration as the six microbeats, and consisted of four snare drum beat sounds with SOAs of 666–333–333–666 ms, presented at 60 dB. Since the training rhythm (snare drum) sounds were presented at a higher intensity than the microbeat (slapstick) sounds, the rhythm masked the microbeat when the two coincided on beats 1, 3, 4, and 5. As a result, the slapstick alone was audible on beats 2 and 6 (the “rest” beats of the rhythm pattern). The training rhythm repeated continuously for the remainder of the 2-min training period (for a total of 60 repetitions). SOAs of the beats in all stimulus component patterns fell within the optimal range (300–800 ms interonset interval) for tempo discrimination (see Baruch & Drake, 1997; Fraisse, 1982).

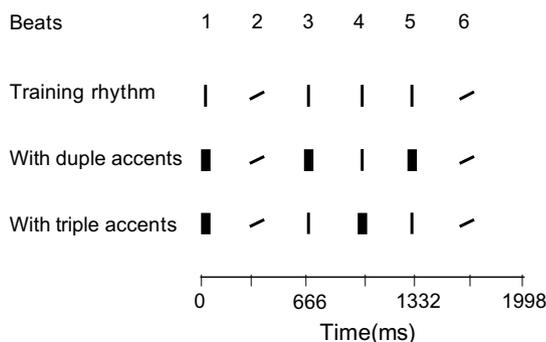


Fig. 1. Stimulus and movement patterns. Beats numbered 1 through 6 equal one musical bar, lasting 1998 ms. Vertical lines represent the snare drum sounds of the rhythm pattern and oblique lines represent time-marking slapstick sounds. The training rhythm had no auditory accents. The training movement occurred in one of two conditions (marked by thick black lines): with duple accents, or with triple accents. Adults heard test stimuli in both forms: in duple form, with auditory accents every second beat, versus triple form, with auditory accents every third beat.

2.1.2.2. Test stimuli. Two test stimuli (Fig. 1) were constructed to be identical to the training rhythm, except that in the test rhythms, accented beats had a relatively high intensity level. This was achieved by maintaining the “strong beats” (the accented snare drum beats) at the same intensity level (60 dB) as the snare drum beats in the training stimulus, while decreasing the intensity level of the “weak beats” (the unaccented snare drum beats) to 55 dB. The *duple rhythm* stimulus subdivided the rhythm pattern into three groups of two beats, with every second beat accented (i.e., BEAT–rest–BEAT–beat–BEAT–rest). The *triple rhythm* stimulus subdivided the rhythmic pattern into two groups of three beats, with every third beat accented (i.e., BEAT–rest–beat–BEAT–beat–rest). Note that both test stimuli were of the pattern *beat–rest–beat–beat–beat–rest*, the difference in metrical interpretation being which beats were accented. The resulting perceptual interpretations sound and feel very different, and adults often do not realize that they are of an identical rhythm pattern.

2.1.3. Apparatus

The auditory stimuli were created using Cakewalk, and recorded as instrumental sounds (i.e., snare drum no. 229, slapstick no. 244) using a Roland 64-Voice Synthesizer Module. Sound files were recorded with Cool Edit 2000 on a personal computer using the AOpen AW-840 4 Channel PCI Sound Card. Stimulus sound files were transferred to a Power Macintosh 7300/180 computer and converted into System 7 sound files for testing. Sounds were presented from a Denon PMA-480R amplifier over Sennheiser headphones to both experimenter and subject. The experiment was run by a custom software program with a custom interface to an experimenter-controlled button box.

2.1.3.1. Seesaw. A seesaw-like bed was custom built of softwood two by fours, with a 12.7 mm diameter steel rod axel, supported by two bearing blocks at its ends. Bungee cords attached each end of the bed to the supporting platform, to add a fixed pulling force at each end. This helped to stabilize the mechanic linkages, provided a more human, joint-like bouncing feel to the rocking motion of the bed, and aided the experimenter in bearing the weight of the subject while performing the rocking motion. The subject was supine on the custom made seesaw both during training and during testing. The subject’s position was adjusted to provide equal distribution of weight over the axel of the seesaw, and to allow for ease of motion. Auditory stimuli during training and test were always presented over headphones.

2.1.4. Procedure

2.1.4.1. Training. The subject was lying on his or her back on the seesaw bed while the experimenter rocked the seesaw on the designated beats. The experimenter’s movement was a gentle bouncing by repeatedly bending at the knees on specified beats. The experimenter’s hands, which pushed

the head of the seesaw down and up repeatedly, remained at waist-level while she bent her knees, so that their movement was aligned with her body movement. Thus the movement experienced by the subject resembled the trajectory of natural body movement. In this fashion, on the downward beats the top of the subject's body moved downwards, while the feet moved upwards, and vice versa on the upward motion. The maximum displacement was approximately three inches up and three inches down from resting position. The experimenter rocked the seesaw throughout the 2-min training phase, for 60 continuous repetitions of the training rhythm stimulus. Subjects were assigned to one of two movement conditions in the training phase (see Fig. 1). In the duple movement condition, rocking occurred on every second beat (beats 1, 3, and 5). In the triple movement condition, rocking occurred on every third beat (beats 1 and 4). In other words, during the training phase the rocking movement provided the accents on the rhythmic strong beats of each subject's respective condition, while the auditory training stimulus was identical in both conditions. Thus the only difference in the training of the two groups of adults was the beats on which they were rocked.

2.1.4.2. Test. Immediately following the training phase the subject remained lying down comfortably on the seesaw bed (which was fixed in resting position), and was given a two-alternative forced-choice task. The experimenter sat across the room, out of sight of the subject, for the remainder of the experiment. The subject listened to the test stimuli over the headphones, while the experimenter used the button box to present eight test trials. Each trial contained a duple and a triple test stimulus, presented in random order. Presentation of the duple and triple rhythms was counterbalanced for trial 1, so that half of the subjects in each condition heard the duple rhythm first, and half heard the triple rhythm first. Subjects were instructed to choose which of the two stimuli was the same as, or most similar to, the sounds they had heard in the training phase. Subjects were never instructed to recall or match to the movement experience; they were only asked to recall the sound.

2.2. Results

We observed a significant effect of movement on stimulus choice. The frequency with which the two groups identified the duple stimulus as familiar at test differed significantly, $t(6) = 8.40$, $p < .0001$, Cohen's $d = 5.95$ (Fig. 2a). Each group thus demonstrated a significant bias to select the stimulus that matched their movement experience; we will henceforth refer to this metrical bias as "performance accuracy", inasmuch as it reflects the successful transfer of the subjects' metrical encoding to their recognition of the matching auditory test stimulus. The two groups did not differ significantly in accuracy; they identified as 'same' the auditory rhythm form that matched their own movement experience with 83% mean performance accu-

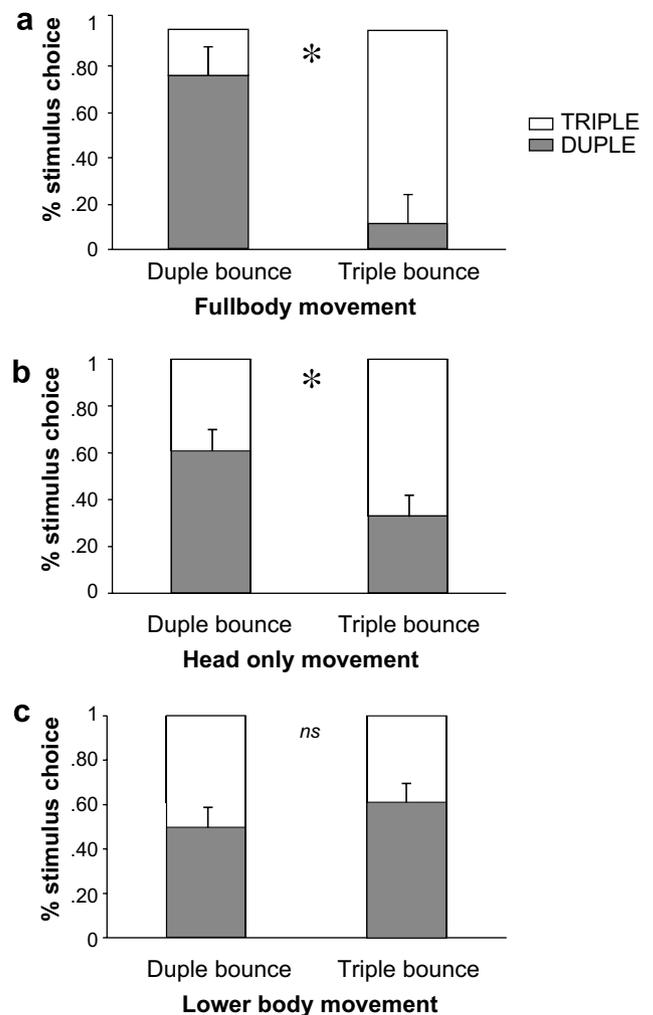


Fig. 2. Results of three experiments. Mean proportion stimulus choice (measure of metrical bias) is represented on the y-axis and type of passive movement training on the x-axis. (a) Experiment 1: Adults who experienced passive body movement identified as 'same' the auditory test stimulus with metrical form matched to their movement experience. (b) Experiment 2: Adults who experienced only head movement also identified the matching auditory test stimulus. (c) Experiment 3: Adults who experienced only movement of the lower body failed to identify the matching auditory test stimulus. Error bars represent the standard error of the mean.

racy. These results were not significantly different from those obtained in the previous study (Phillips-Silver & Trainor, 2007) which employed active, self-generated movement on the part of the participant, $t(14) = -.30$, $t(2-tailed) > .05$.

Seven subjects reported having musical training, and there was a significant positive correlation between the number of years of musical training (range 0–15) and task performance, $r = .77$, $N = 8$, $p(2-tailed) = .03$. While in the four experiments of our previous set of studies (Phillips-Silver & Trainor, 2007) there were no significant correlations between music training and task performance, in the present experiment musical experience was reported by a greater proportion of subjects, and with a greater range, which may explain this significant correlation. Because all

8 subjects reported some kind of recreational music activity, and recreational activity was a dichotic variable, no correlations could be performed with this variable. There was no significant correlation between test accuracy and recreational dance.

The conservation of the multisensory effect under the passive motion condition with reduced motor planning, tactile, and proprioceptive information suggests that vestibular input might be critical. The goal of Experiment 2 was to further isolate the activation of the vestibular system by applying motion to the head only.

3. Experiment 2

3.1. Method

3.1.1. Participants

This study included 16 healthy university undergraduate students (aged 19–40 years, mean 22.9 years). Musical training ranged from 0 to 14 years (mean 5.2 years). Fourteen of 16 reported some recreational music activity; 8 of 16 reported recreational dancing.

3.1.2. Stimuli, apparatus, and procedure

The stimuli, apparatus, and procedure were identical to those of Experiment 1 with the following exceptions. Participants sat in a chair at the head of the seesaw, with their back and shoulders supported by the chair back. The chair was tilted back at approximately a 45-degree angle, and the participant's head rested on a pillow at the end of the bed. Thus the head and spine were aligned, allowing for flexion of the neck muscles (while the torso was held stable by the chair back) when the end of the seesaw supporting the head dipped downwards on the designated beats. The experimenter lowered and raised the seesaw as in Experiment 1. The participant's head rocked down and up with approximately six-inch total displacement, equivalent to the displacement of the head and feet of participants in Experiment 1.

3.2. Results

We observed a modest but significant effect of head movement on stimulus choice. The frequency with which the two groups identified the duple stimulus as familiar at test differed significantly, as measured by an independent samples *t*-test, $t(14) = 1.57$, $p = .05$, Cohen's $d = 0.78$. The two groups did not differ in accuracy; participants chose their matching stimulus with 66% mean accuracy, which was significantly above chance performance, $t(15) = 1.84$, $p < .05$ (Fig. 2b). There were no significant correlations between test accuracy and number of years of musical training or recreational music activity. There was a significant correlation between test accuracy and recreational dance, $r = .43$, $N = 16$, $p(1\text{-tailed}) = .05$.

The accuracy of participants in the head-only motion condition was significantly lower than the accuracy of

participants in the full-body motion condition of Experiment 1, $t(22) = 1.39$, $p = .05$, Cohen's $d = 0.69$. Several factors may have contributed to the reduced magnitude of the effect with head movement alone. First, there may be a reduction in the vestibular stimulation with head-only motion as compared with whole body motion. When the whole body is rotated a vestibulocollic reflex, which functions to stabilize the head in space, produces a compensatory head-on-trunk movement (Gdowski & McCrea, 1999). This reflexive response affects vestibular processing; specifically, there are secondary vestibular neurons that respond selectively to the head-on-trunk motion when both are moving (Gdowski & McCrea, 1999). Thus it is possible that information from these secondary vestibular neurons was available in our whole-body condition, but inhibited if not absent in the head-only condition, allowing for a more accurate neuronal assessment of head and body movement in the former case. Second, the movement of the head in the gravitational field, and hence the specific stimulation of the vestibular organs, is somewhat different in Experiments 1 and 2. It is possible that some head movements are more effective than others at inducing the effect of movement on auditory rhythm perception. Third, it is possible that the subject lying on the seesaw and experiencing whole body rocking feels more comfortable than the subject leaning back in the chair and having their head moved from the neck. In everyday life, we may have our whole body hugged and rocked more often than we have our heads moved from the neck.

Finally, the effect of movement on auditory metrical interpretation may not come from the vestibular system alone. For example, proprioceptive information might contribute to the interaction between movement and audition, and the amount of overall proprioceptive information is likely reduced when only the head and neck are moved compared to when the full body is moved. In order to test the contribution of aspects of movement other than vestibular, in Experiment 3 we tested whether the effect would remain when subjects' legs and feet were moved but their head was not (vestibular information absent).

4. Experiment 3

4.1. Method

4.1.1. Participants

This study included 16 healthy university undergraduate students (aged 19–40 years, mean 22 years). Musical training ranged from 0 to 16 years (mean 8.1 years). Twelve of 16 reported some recreational music activity; 3 of 16 reported recreational dancing.

4.1.2. Stimuli, apparatus, and procedure

The stimuli, apparatus, and procedure were identical to those of Experiment 1 with the following exceptions. Experiment 3 employed the same motion as in Experiments 1 and 2, but applied to the feet and legs only. Participants

were supine on the floor at the head of the seesaw, and their feet rested on the end of the bed so that their legs were bent at approximately a 90-deg angle. The experimenter performed the bouncing motion in the same manner as Experiment 1, lowering and raising the seesaw by plus or minus three inches. Consequently, the participants' feet bounced down and up (and their legs bent or extended further), with approximately six-inch total displacement of the feet.

4.2. Results

Experiment 3 revealed no significant effect of lower body (feet and legs only) movement on stimulus choice (Fig. 2c). The frequency with which the two groups identified the duple stimulus as familiar at test did not differ significantly, $t(14) = -.61$, $p > .05$, Cohen's $d = 0.31$. The accuracy with which each group chose the matching stimulus was not different from chance. There were no significant correlations between test accuracy and number of years of musical training, recreational music activity, or recreational dance.

The percent correct score from Experiment 3 was also significantly lower than that of participants in Experiment 2, $t(30) = 1.73$, $p < .05$, Cohen's $d = 0.61$. Since vestibular input is present when the head is moving (Experiments 1 and 2), but greatly reduced, if not eliminated, when the head is still (Experiment 3), we propose that vestibular stimulation may be a critical component to the observed effect of movement on auditory rhythm perception.

In summary we find a gradient of diminishing effect on rhythm perception, from whole body motion, to head motion alone, to lower body motion alone. It appears that each of these conditions progressively eliminates components of the movement input that affect auditory perception. Critically, vestibular information alone appears to be strong enough to guide the metrical interpretation of ambiguous rhythm patterns.

5. Discussion

The main goal of these experiments was to determine whether vestibular input contributes to the influence of movement on the metrical interpretation of an auditory rhythm pattern. In Experiment 1, participants were rocked on either every second or every third beat of an ambiguous rhythm pattern while lying passively on a seesaw, which greatly reduced motor, tactile, and proprioceptive inputs (e.g., Kandel, 2000) compared to the case of previous studies in which participants stood and bounced their bodies in synchrony with an experimenter (Phillips-Silver & Trainor, 2007).

Nevertheless, the influence of movement on disambiguating the metrical structure of the auditory rhythm pattern was as strong with the passive as with the active movement for adults, suggesting the potential involvement of the vestibular system. With passive movement of the kind we employed here, it is difficult to entirely eliminate the

influence of motor, tactile, and proprioceptive input. For example, although the vestibular system does not differentiate passive from active head movement (Cullen & Minor, 2002), proprioceptive input about head orientation with respect to the trunk is provided by the tension of the muscle spindles of the neck, which changes as the neck moves (Lewald, Karnath, & Ehrenstein, 1999; Snyder, Grieve, Brotchie, & Andersen, 1998). However, the influence of vestibular input can be examined by comparing passive movement that either involves the head or does not involve the head. In Experiments 2 and 3 we compared passive movement of the head to passive movement of the legs and feet, and found that only when the head was moved did participants show a metrical bias. We therefore have strong evidence that the vestibular input contributes to the interaction between movement and auditory systems in the perception of metrical structure. This conclusion is strengthened by recent converging evidence that direct stimulation of the vestibular nerve can influence the perception of ambiguous metrical patterns (Trainor, Gao, Lei, Lehtovaara, & Harris, in press).

Exactly where auditory and vestibular information is integrated in the nervous system remains a mystery, although many projections from brainstem to cortex have polymodal components (Linke & Schwegler, 2000). There are many potential sites for direct and indirect interaction, and the process likely involves complex networks of cortical and subcortical areas. Both the hearing and vestibular end organs are located in the inner ear, but although very loud levels of sound, such as those found at rock concerts, may directly stimulate the semi-circular canals (Todd & Cody, 2000), under normal listening conditions the information that they collect is funneled through the separate auditory and vestibular nerve channels. At the subcortical level, visual–vestibular pathways have been well studied (Nolte, 2002; Schlack, Hoffman, & Bremmer, 2002), but evidence for auditory–vestibular convergence is scant. Nevertheless, both the auditory and vestibular systems develop early in utero and are functioning in the third trimester (Romand, 1992), and a clear interaction between movement and auditory rhythm perception has been shown behaviorally in 7-month-old infants (Phillips-Silver & Trainor, 2005). If the vestibular–auditory effects are indeed cortically mediated, no multisensory effect would be expected in infants 2 months of age or younger as the auditory cortex is not mature enough at this stage to support complex processing (Moore & Guan, 2002). One recent study suggests that auditory and vestibular information indeed converge as early as the dorsal cochlear nuclei (DCN) (Oertel & Young, 2004). Shiroyama, Kayahara, Yasui, Nomura, and Nakano (1999) show that there are vestibular inputs to auditory thalamic nuclei that project to auditory cortical areas, in particular, through the medial geniculate body. The cerebellum is another structure of interest as it is known to be involved in auditory rhythm processing (Griffiths, 2003; Parsons, 2003; Penhune, Zatorre, & Evans, 1998; but see Molinari et al., 2005)

and it receives major vestibular as well as auditory input (e.g., Suzuki & Keller, 1982).

At the cortical level, there is a growing body of evidence for the integration of multisensory cues, including visual, tactile, vestibular, and auditory signals, in coding spatial and motion information in human and non-human primates (Bremmer, 2005). For example, moving towards an object requires a retinal image of the object as well as information about body coordinates as it moves through space, and the transformation of the visual input into a body or world frame of reference is attributed to neurons in the posterior parietal cortex (PPC) (for a review, see Bremmer, 2005). The same region has recently been investigated for the integration of vestibular and acoustic information in tasks of sound localization and space perception, and has been proposed as part of a dorsal (“where”) stream in auditory cortical processing (Lewald, Foltys, & Töpper, 2002).

A subregion of the PPC, the ventral intraparietal area (VIP), contains polymodal neurons that respond to moving stimuli, many of which are proposed to encode inputs from different sensory modalities into a head-centered frame of reference (Bremmer, Schlack, Duhamel, Graf, & Fink, 2001). Area VIP has substantial inputs from auditory areas; as well, it forms part of a cortical vestibular network (Lewis & Van Essen, 2000). Single cell responses in the macaque area VIP have been elicited by auditory stimulation (Schlack, Sterbing-D’Angelo, Hartung, Hoffmann, & Bremmer, 2005) as well as by visual, vestibular (self-motion) and somatosensory stimulation (Bremmer, Schlack et al., 2001; Colby, Duhamel, & Goldberg, 1993; Schlack et al., 2002). This region is specialized to combine sensory signals such as visual (i.e., optical flow, eccentric eye position) and vestibular information in guiding self-motion through the environment (e.g., Bremmer, Klam, Duhamel, Ben Hamed, & Graf, 2002; Schlack et al., 2002; Telford, Howard, & Ohmi, 1995).

Recent studies have revealed an effect of vestibular afferent information on the perception of sound location during movement of the head and the whole body (Lewald & Karnath, 2002). By tilting the whole body of adult participants, Lewald and Karnath (2002) demonstrated a direct influence of otolith vestibular information on neural processing of auditory spatial cues (Lewald & Karnath, 2002). Repetitive transcranial magnetic stimulation (rTMS) interrupts neural processing in the PPC and alters perception of sound location without affecting binaural processing acuity, supporting the role of the human PPC in integration of auditory and vestibular cues such as in spatial hearing (Lewald et al., 2002). Lewald and colleagues suggest that the PPC is responsible for neural coordinate transformations, that is, from originally head-centered to body- or world-centered auditory space, which may be required for perceptual stability of auditory and multisensory space during self-motion (Lewald, Wienemann, & Boroojerdi, 2004; Their & Karnath, 1997). The integration of auditory with vestibular and

proprioceptive inputs can be crucial to the ability to navigate the environment, as can be seen in early blind individuals. In the absence of vision, blind individuals use audiomotor feedback to calibrate auditory space (Lewald, 2002). Vision is not necessary for space perception, and vision loss is not compensated by a general improvement in auditory acuity or in distance estimation, but rather by enhanced processing of proprioceptive and vestibular information with auditory spatial information (Lewald, 2002; Loomis, Klatzky, & Golledge, 2001; Ruddle & Lessels, 2006). Consistent with this view, Weeks et al. (2000) demonstrated enhanced PPC activation in sound localization in blind individuals, providing further support for the role of this brain region in the integration of auditory and vestibular information.

The view of the PPC as an integrator of acoustic and vestibular cues, together with evidence of area VIP as a site of multimodal neurons that code for spatial perception and self-motion, offer an account of the auditory–vestibular connections that may underlie our findings of multisensory interactions between movement and the perception of auditory rhythm. However, other areas must be considered as well. It is known that the parietoinsular vestibular cortex is responsive to vestibular, optokinetic, and proprioceptive stimulation from muscle and joint receptors, such as in the neck (Guldin, Akbarian, & Grüsser, 1992). Another major site of interest is the prefrontal cortex, which receives multisensory inputs including vestibular and visual inputs, and which participates in the vestibular representation of the body’s orientation and displacement in space (Israël, Rivaud, Gaymard, Berthoz, & Pierrot-Deseilligny, 1995). With respect to auditory–motor integration in music, recent neuroimaging evidence implicates the superior temporal gyrus and the dorsal premotor cortex in the synchronization of movement to musical rhythms, structures for which direct anatomical connections have been shown in non-human primates (Chen et al., 2006). Other regions that potentially contribute to such auditory–motor interactions and that share direct connections with each other include posterior auditory regions and the ventral premotor cortex, which responds to auditory stimuli (Bremmer et al., 2001; Mesulam & Mufson, 1982; Seltzer & Pandya, 1989), as well as the insula, which is a multimodal structure involved in the temporal integration of sensory stimuli and detection of stimulus synchrony, and participates in auditory–motor interactions between posterior auditory regions and dorsal premotor cortex (Bushara, Grafman, & Hallett, 2001; Calvert, 2001; Lux, Marshall, Ritzl, Zilles, & Fink, 2003).

While more research is evidently needed in order to understand the neural underpinnings of how the vestibular and auditory systems work together, it is clear that musical rhythm patterns elicit movement, that movement of the body can influence auditory perception of the metrical structure of rhythm, and that vestibular and auditory information are integrated in perception.

Acknowledgments

This research was supported by a grant to L.J.T. from the Natural Sciences and Engineering Research Council of Canada. We thank Miroslav Cika for custom building the experimental apparatus, and members of the Auditory Development Laboratory at McMaster University for discussions on theoretical content and methodology.

References

- Baruch, C., & Drake, C. (1997). Tempo discrimination in infants. *Infant Behaviour and Development*, 20(4), 573–577.
- Bremmer, F. (2005). Navigation in space—The role of the macaque ventral intraparietal area. *The Journal of Physiology*, 566, 29–35.
- Bremmer, F., Klam, F., Duhamel, J.-R., Ben Hammed, S., & Graf, W. (2002). Visual–vestibular interactive responses in the macaque ventral intraparietal area (VIP). *Neuroreport*, 10, 873–878.
- Bremmer, F., Schlack, A., Duhamel, J.-R., Graf, W., & Fink, G. R. (2001). Space coding in primate posterior cortex. *Neuroimage*, 14, S46–S51.
- Bremmer, F., Schlack, A., Shah, N. J., Zafiris, O., Kubischik, M., Hoffmann, K., et al. (2001). Polymodal motion processing in posterior parietal and premotor cortex: A human fMRI study strongly implies equivalencies between humans and monkeys. *Neuron*, 29, 287–296.
- Brown, W. M., Lee, C., Grochow, K., Jacobson, A., Liu, C. K., Popovic, Z., et al. (2005). Dance reveals symmetry especially in young men. *Nature*, 438(7071), 1148–1150.
- Bushara, K. O., Grafman, J., & Hallett, M. (2001). Neural correlates of auditory–visual stimulus onset asynchrony detection. *Journal of Neuroscience*, 21, 300–304.
- Calvert, G. A. (2001). Crossmodal processing in the human brain: Insights from functional neuroimaging studies. *Cerebral Cortex*, 11, 1110–1123.
- Chen, J. L., Zatorre, R. J., & Penhune, V. B. (2006). Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms. *Neuroimage*, 32, 1771–1781.
- Clark, D. L., Kreutzberg, J. R., & Chee, F. K. (1977). Vestibular stimulation influence on motor development in infants. *Science*, 196(4295), 1228–1229.
- Clarke, E. F. (1993). Imitating and evaluating real and transformed musical performances. *Music Perception*, 10, 317–341.
- Clarke, E. F. (1999). Rhythm and timing in music. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 473–500). NY: Academic Press.
- Colby, C. L., Duhamel, J. R., & Goldberg, M. E. (1993). Ventral intraparietal area of the macaque: Anatomic location and visual response properties. *Journal of Neurophysiology*, 79, 126–136.
- Cullen, K. E., & Minor, L. B. (2002). Semicircular canal afferents similarly encode active and passive head-on-body rotations: Implications for the role of vestibular efference. *Journal of Neuroscience*, 22:RC226, 1–7.
- Cullen, K. E., & Roy, J. E. (2004). Signal processing in the vestibular system during active versus passive head movements. *Journal of Neurophysiology*, 91(5), 1919–1933.
- Driver, A. (1936). *Music and movement*. London: Oxford University Press.
- Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149–180). New York: Academic Press.
- Gdowski, G. T., & McCrea, R. A. (1999). Integration of vestibular and head movement signals in the vestibular nuclei during whole-body rotation. *Journal of Neurophysiology*, 82, 436–449.
- Goldberg, J. M., & Fernandez, C. (1980). Efferent vestibular system in the squirrel monkey: Anatomical location and influence on afferent activity. *Journal of Neurophysiology*, 43, 986–1025.
- Griffiths, T. (2003). The neural processing of complex sounds. In I. Peretz & R. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 167–177). New York: Oxford University Press.
- Guldin, W. O., Akbarian, S., & Grüsser, O. J. (1992). Cortico-cortical connections and cytoarchitectonics of the primate vestibular cortex: A study in squirrel monkeys (*Saimiri sciureus*). *Journal of Comparative Neurology*, 326, 375–401.
- Hauelsen, J., & Knosche, T. R. (2001). Involuntary motor activity in pianists evoked by music perception. *Journal of Cognitive Neuroscience*, 13, 786–792.
- Israël, I. R., Rivaud, S., Gaymard, B., Berthoz, A., & Pierrot-Deseilligny, C. (1995). Cortical control of vestibular-guided saccades in man. *Brain*, 118, 1169–1183.
- Jacques-Dalcroze, É. (1920). *Method of eurhythmics: Rhythmic movement* (Vol. 1). London: Novello & Co.
- Juntunen, M.-L., & Hyvönen, L. (2004). Embodiment in musical knowing: How body movement facilitates learning with Dalcroze Eurhythmics. *British Journal of Music Education*, 21(2), 199–214.
- Kandel, E. R. (2000). Nerve cells and behavior. In E. R. Kandel, J. H. Schwartz, & T. M. Jessel (Eds.), *Principles of neural science* (4th ed., pp. 19–35). New York: McGraw Hill.
- Klam, F., & Graf, W. (2003). Vestibular signals of posterior parietal cortex neurons during active and passive head movements in macaque monkeys. *Annals of the New York Academy of Science*, 1004, 271–282.
- Klinke, R., & Schmidt, C. L. (1970). Efferent influence on the vestibular organ during active movements of the body. *Pflugers Archiv* 318, 325, as cited In E. R. Kandel, J. H. Schwartz, & T. M. Jessel (Eds.), *Principles of neural science* (4th ed., pp. 19–35). New York: McGraw Hill.
- Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: Audiomotor recognition network while listening to newly acquired actions. *Journal of Neuroscience*, 27, 308–314.
- Lewald, J. (2002). Opposing effects of head position on sound localization in blind and sighted human subjects. *European Journal of Neuroscience*, 15, 1219–1224.
- Lewald, J., Foltys, H., & Töpper, R. (2002). Role of the posterior parietal cortex in spatial hearing. *The Journal of Neuroscience*, 22:RC207, 1–5.
- Lewald, J., & Karnath, H.-O. (2002). The effect of whole-body tilt on sound localization. *European Journal of Neuroscience*, 16, 761–766.
- Lewald, J., Karnath, H.-O., & Ehrenstein, W. H. (1999). Neck-proprioceptive influence on auditory lateralization. *Experimental Brain Research*, 125, 389–396.
- Lewald, J., Wienemann, M., & Boroojerdi, B. (2004). Shift in sound localization induced by rTMS of the posterior parietal lobe. *Neuropsychologia*, 42, 1598–1607.
- Lewis, J., & Van Essen, D. C. (2000). Corticocortical connections of visual, sensorimotor and multimodal processing areas in the parietal lobe of the macaque monkey. *Journal of Computational Neurology*, 428, 112–137.
- Linke, R., & Schwegler, H. (2000). Convergent and complementary projections of the caudal paralaminar thalamic nuclei to rat temporal and insular cortex. *Cerebral Cortex*, 10, 753–771.
- Loomis, J. M., Klatzky, R. L., & Golledge, R. G. (2001). Navigating without vision: Basic and applied research. *Optometry and Vision Science*, 78(5), 282–289.
- Lux, S., Marshall, J. C., Ritzl, A., Zilles, K., & Fink, G. R. (2003). Neural mechanisms associated with attention to temporal synchrony versus spatial orientation: An fMRI study. *Neuroimage*, 20, S58–S65.
- Mesulam, M. M., & Mufson, E. J. (1982). Insula of the old world monkey: III. Efferent cortical output and comments on function. *Journal of Comparative Neurology*, 212, 38–52.
- Molinari, M., Leggio, M. G., Filippini, V., Gioia, M. C., Cerasa, A., & Thaut, M. (2005). Sensorimotor transduction of time information is preserved in subjects with cerebellar damage. *Brain Research Bulletin*, 67(6), 448–458.
- Moore, J. K., & Guan, Y. L. (2002). Cytoarchitectural and axonal maturation in human auditory cortex. *Journal of the Association for Research in Otolaryngology*, 2, 297–311.
- Excerpts from hearing and balance: The eighth cranial nerve, (2002). In J. Nolte (Ed.) *The human brain: An introduction to its functional anatomy*, (5th ed.). Mosby, Inc.
- Oertel, D., & Young, E. D. (2004). What's a cerebellar circuit doing in the auditory system? *Trends in Neuroscience*, 27, 104–110.

- Palmer, C. (1997). Music performance. *Annual Reviews of Psychology*, 48, 115–138.
- Parsons, L. M. (2003). Exploring the functional neuroanatomy of music performance, perception and comprehension. In I. Peretz & R. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 247–268). New York: Oxford University Press.
- Penhune, V. B., Zatorre, R. J., & Evans, A. E. (1998). Cerebellar contributions to motor timing: A PET study of auditory and visual rhythm reproduction. *Journal of Cognitive Neuroscience*, 10, 752–765.
- Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science*, 308, 1430.
- Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition*, 105, 533–546.
- Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin and Review*, 12, 969–992.
- Repp, B. H., & Doggett, R. (2007). Tapping to a very slow beat: A comparison of musicians and nonmusicians. *Music Perception*, 24(4), 367–376.
- Romand, R. (Ed.). (1992). *Development of auditory and vestibular systems*. Amsterdam, New York: Elsevier.
- Ruddle, R. A., & Lessels, S. (2006). For efficient navigational search, humans require full physical movement, but not a rich visual scene. *Psychological Science*, 17(6), 460–465.
- Schlack, A., Hoffman, K. P., & Bremner, F. (2002). Interaction of linear vestibular and visual stimulation in the macaque ventral intraparietal area (VIP). *European Journal of Neuroscience*, 16, 1877–1886.
- Schlack, A., Sterbing-D'Angelo, S. J., Hartung, K., Hoffmann, K.-P., & Bremner, F. (2005). Multisensory space representations in the macaque ventral intraparietal area. *The Journal of Neuroscience*, 25(18), 4616–4625.
- Seltzer, B., & Pandya, D. N. (1989). Frontal lobe connections of the superior temporal sulcus in the rhesus monkey. *Journal of Comparative Neurology*, 281, 97–113.
- Shahidullah, S., & Hepper, P. G. (1994). Frequency discrimination by the fetus. *Early Human Development*, 36, 13–26.
- Shiroyama, T., Kayahara, T., Yasui, Y., Nomura, J., & Nakano, D. (1999). Projections of the vestibular nuclei to the thalamus in the rat: A phaseolus vulgaris leucoagglutinin study. *Journal of Comparative Neurology*, 407, 318–332.
- Shove, P., & Repp, B. H. (1995). Musical motion and performance: Theoretical and empirical perspectives. In J. Rink (Ed.), *The practice of performance: Studies in musical interpretation* (pp. 55–83). New York: Cambridge University Press.
- Snyder, L. H., Grieve, K. L., Brotchie, P., & Andersen, R. A. (1998). Separate body- and world-referenced representations of visual space in parietal cortex. *Nature*, 394(6696), 887–891.
- Suzuki, D. A., & Keller, E. L. (1982). Vestibular signals in the posterior vermis of the alert monkey cerebellum. *Experimental Brain Research*, 47, 145–147.
- Telford, L., Howard, I. P., & Ohmi, M. (1995). Heading judgments during active and passive self-motion. *Experimental Brain Research*, 104, 502–510.
- Their, P., & Karnath, H.-O. (Eds.). (1997). *Parietal lobe contributions to orientation in 3D space*. Heidelberg: Springer.
- Thompson, W. F., Graham, P., & Russo, F. A. (2005). Seeing music performance: Visual influences on perception and experience. *Semiotica*, 156, 177–201.
- Todd, N. P. M. (1995). The kinematics of musical expression. *Journal of the Acoustical Society of America*, 97(3), 1940–1949.
- Todd, N. P. M. (1999). Motion in music: A neurobiological perspective. *Music Perception*, 17(1), 115–126.
- Todd, N. P. M., & Cody, F. W. (2000). Vestibular responses to loud dance music: A physiological basis of the “rock and roll threshold”? *Journal of the Acoustical Society of America*, 107(1), 496–500.
- Todd, N. P., Cousins, R., & Lee, C. S. (2007). The contribution of anthropomorphic factors to individual differences in the perception of rhythm. *Empirical Musicology Review*, 2(1), 1–13.
- Trainor, L. J. (2007). Do preferred beat rate and entrainment to the beat have a common origin in music? *Empirical Music Review*, 2, 17–20.
- Trainor, L. J., Gao, X., Lei, J., Lehtovaara, K., & Harris, L. R. (in press). The primal role of the vestibular system in determining musical rhythm. *Cortex*.
- Trehub, S. (2003). The developmental origins of musicality. *Nature Neuroscience*, 6(7), 669–673.
- Weeks, R., Horwitz, B., Aziz-Sultan, A., Tian, B., Wessinger, M., Cohen, L. G., et al. (2000). A Positron Emission Tomographic study of auditory localization in the congenitally blind. *Journal of Neuroscience*, 20, 2664–2672.
- 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.