

Provided for non-commercial research and education use.  
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



ELSEVIER

available at [www.sciencedirect.com](http://www.sciencedirect.com)journal homepage: [www.elsevier.com/locate/cortex](http://www.elsevier.com/locate/cortex)**Special issue: Research report**

# The primal role of the vestibular system in determining musical rhythm

Laurel J. Trainor<sup>a,b,\*</sup>, Xiaoqing Gao<sup>a</sup>, Jing-jiang Lei<sup>a</sup>, Karen Lehtovaara<sup>a</sup> and Laurence R. Harris<sup>c</sup>

<sup>a</sup>Department of Psychology, Neuroscience & Behaviour and the McMaster Institute for Music and the Mind, McMaster University, Hamilton, Ontario, Canada

<sup>b</sup>Rotman Research Institute, Baycrest Hospital, Toronto, Ontario, Canada

<sup>c</sup>Centre for Vision Research, York University, Toronto, Ontario, Canada

## ARTICLE INFO

### Article history:

Received 19 January 2007

Reviewed 30 May 2007

Revised 20 June 2007

Accepted 4 October 2007

Published online 7 November 2008

### Keywords:

Vestibular stimulus

Galvanic stimulation

Musical rhythm

Vestibulo-auditory interaction

## ABSTRACT

Previous studies have indicated that physical movement on either every second or on every third beat of an unaccented auditory rhythm pattern can disambiguate whether it is perceived in duple time as a march or in triple time as a waltz. Here we demonstrate that this disambiguation can also be accomplished by direct galvanic stimulation of the vestibular system. The galvanically induced sensation, without any actual movement, that the head moved from side to side on either every second or on every third beat of the ambiguous auditory rhythm pattern strongly biased whether adults perceived it as being in duple or in triple time. These results imply that the vestibular system plays a primal role in the perception of musical rhythm.

© 2008 Elsevier Srl. All rights reserved.

## 1. Introduction

Physical movement and the perception of musical rhythm seem to be closely bound. Historically, music and dance have evolved together, and across all cultures, people move in synchrony to musical rhythms (Arom, 1991; Clarke, 1999; Cross, 2003; Molinari et al., 2003; Todd, 1995; Wallin et al., 2000). Body movement is commonly used in music education in order to teach rhythmic patterns, further suggesting a link between movement and the perception of rhythm. For example, in Kindermusik classes infants are passively moved

to music with different rhythmic patterns (Cutietta, 2001) and more advanced students in Dalcroze Eurhythmics classes learn to internalize rhythmic patterns through physical movement in time and space (Jaques-Dalcroze, 1920; Juntunen and Hyvönen, 2004). Adults and children can readily tap to the strong beats of a rhythmic pattern, regardless of musical training (Drake et al., 2000) and tapping behaviour reflects their auditory representations of the rhythm pattern (Repp, 2005, review). Clearly there is a deep connection between movement and rhythm but how might this have come about?

\* Corresponding author. Department of Psychology, Neuroscience & Behaviour and the McMaster Institute for Music and the Mind, McMaster University, Hamilton, ON L8S 4K1, Canada.

E-mail addresses: [ljt@mcmaster.ca](mailto:ljt@mcmaster.ca) (L.J. Trainor), [harris@yorku.ca](mailto:harris@yorku.ca) (L.R. Harris).

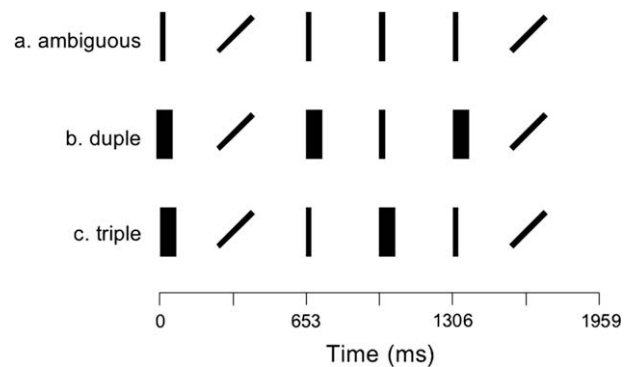
0010-9452/\$ – see front matter © 2008 Elsevier Srl. All rights reserved.

doi:10.1016/j.cortex.2007.10.014

Perhaps the most common form of movement with a tempo is walking. That locomotion involves regularly spaced movements of the arms, legs, and the head has been observed for a long time. Curiously, however, the tempo range of locomotion movements turns out to correspond closely to the range over which a beat can be felt – pulse separations of between about 300 and 900 msec (e.g., Fraise, 1982; Clarke, 1999). At tempos slower than this, auditory events are not connected into a perceptual pattern and at faster tempos, events cannot be sufficiently individuated. Furthermore an individual's preferred auditory beat rate correlates with anthropometric variables such as height, leg length, shoulder breadth, and body weight (Todd et al., 2007) that are in turn related to an individual's locomotion tempo. These correlations provide circumstantial evidence for a link between the cues generated by locomotion and the auditory perception of tempo. But musical rhythm can be much more complex than the simple repetitive frequency of walking.

One of the remarkable abilities of humans is that we can entrain and feel musical beats at several different tempos. Although many non-human animals engage in rhythmic behaviours (e.g., apes will drum and certain birds will peck rhythmically), only a few species entrain their movements to different externally defined tempos (e.g., cockatoos, Patel et al., 2008). Indeed, it appears that only species that engage in vocal learning exhibit rhythmic entraining (Schachner et al., 2008). Being able to produce different tempos synchronized with other players is a likely prerequisite for human musical behaviour (Trainor, 2007). Humans can extract a complex metrical hierarchy from a rhythm pattern beyond that directly present in the stimulus (e.g., Lerdahl and Jackendoff, 1983). For example, the tactus level (tempo at which you would tap your foot) of a rhythm pattern is readily extracted, even from patterns that may not always contain sound events on the beats that are perceived as strong (i.e., accented). A tactus level can also be derived when the tempo is not steady or even if the rhythm contains accented “off-beat” sound events, as in the syncopations of jazz. Furthermore, the perceived metrical structure is hierarchical. For example, the beats of the tactus can be subdivided into groups of two sub-beats (as in a march) or three sub-beats (as in a waltz), producing a fast tempo that can be perceived at the same time as the ongoing slower tactus. Tempos can also be slowed down by grouping. For example, every second or every third tactus-level beat may be more strongly accented, producing a slower rhythm.

Metrical structure is normally derived from the pattern of physically accented beats – e.g., sound events that are longer, louder, and/or higher in pitch. However, physical movement can also influence the metrical interpretation of an ambiguous rhythm pattern (Phillips-Silver and Trainor, 2005, 2007). Just as ambiguous visual stimuli can be created that can be perceived in two distinct ways, such as the famous Ruben drawing of a vase, the edges of which can alternatively be perceived as two inward-looking faces, Phillips-Silver and Trainor created an ambiguous auditory stimulus. It consisted of a repeating six-beat drum pattern with no accented notes (see Fig. 1). If every second beat was “heard” as strong, the rhythm took on a march-like quality (duple meter) whereas if every third beat was “heard” as strong, the rhythm took on a waltz-like quality (triple meter). The two interpretations of the pattern sounded



**Fig. 1 – The auditory rhythm. Thin vertical lines represent unaccented snare drum beats. Thick vertical lines represent accented snare drum beats. Oblique lines represent rests. The top line represents the ambiguous rhythm used during familiarization. The other two lines represent the rhythms used during the test phase. The middle line represents the rhythm with auditory accents in duple (march) time and the bottom line the rhythm with auditory accents in triple (waltz) time.**

completely different, just as the two interpretations of the Ruben figure look completely different. And just as only one interpretation of an ambiguous visual pattern can be perceived at any one time (either a vase or two faces), only one interpretation of the ambiguous auditory rhythm could be perceived at any one time.

A physical bouncing movement on either every second beat or on every third beat of the ambiguous auditory rhythm pattern biased which interpretation was perceived in both infant (Phillips-Silver and Trainor, 2005) and adult subjects (Phillips-Silver and Trainor, 2007). Visual information was not necessary for the effect as the results were similar when the subjects were blindfolded. On the other hand, physical movement of the subject was critical as passively observing the experimenter moving did not bias whether the ambiguous rhythm sounded like a march or a waltz in either infants or adults. These studies demonstrated the primal role of movement in determining musical rhythm but could not indicate which aspect(s) of movement were critical in determining the perceived metrical interpretation.

What correlate of movement might establish the correlation between physical movement and auditory rhythm? Candidates include motor planning (efferent copy of motor signals), sensory feedback arising from the consequences of the movement such as tactile or visual information, or sensory feedback directly about the movement itself arising from proprioceptive sources. A major source of proprioceptive information concerning whole body movement comes from the vestibular system. Passive movement of the legs on either every second or third beat of an ambiguous auditory rhythm did not bias whether adults perceived a duple or triple metrical structure, but passive movement of the head on these beats did (Phillips-Silver and Trainor, 2008). Given that both the legs and the head are strongly involved in locomotion, sensorimotor planning, tactile input, and proprioceptive input, but only movement of the head generates a vestibular

signal, these results suggest that vestibular input may be crucial to the multisensory interaction between movement and auditory rhythm. The hypothesis of vestibular involvement is consistent with the early emergence of movement/auditory interactions in infancy and the early maturation of the vestibular system, as illustrated by the delight with which premobile infants greet bouncing, rocking, and swooping stimulation.

In this paper, we dissociate the vestibular signal from aspects of movement by stimulating the vestibular system directly in the absence of physical movement using galvanic stimulation (see Buys, 1909; Goldberg et al., 1984; Mars et al., 2005; Wardman et al., 2003; Zink et al., 1997). Subjects listened to an ambiguous rhythm pattern while we stimulated the vestibular nerve in such a way as to create the sensation of a side-to-side movement of the head, timed to occur on every second beat of the auditory rhythm for half of the subjects and on every third beat for the other half. We provide evidence that a vestibular signal alone is sufficient to bias the metrical interpretation of an auditory rhythm pattern.

## 2. Method

### 2.1. Participants

Participants were 23 (6 male, 17 female) university undergraduate students (aged 18–25 years, mean age = 19 years) enrolled in an introductory psychology course who received course credit for participation. All of the participants had no known hearing deficits and were unaware of the purpose of the study. Half of the participants were randomly assigned to the experimental condition and half to the control condition. Two additional subjects were excluded, one due to equipment failure and one for producing data more than 3 standard deviations from the mean. The experimental group was comprised of 11 subjects; the control group was comprised of 12 subjects. Subjects had between 0 to 14 years of music lessons (mean = 4.3 years for the experimental group and 4.1 years for the control group). Procedures were approved by the McMaster University Ethics Board and subjects gave written consent to participate.

### 2.2. Stimuli

#### 2.2.1. Auditory stimuli

Both the familiarization and test stimuli were similar to those of Phillips-Silver and Trainor (2005) and were presented in a sound-attenuating chamber over a noise floor of 25 dB(A). The vestibular stimulation was presented with pulses every 25.1 msec, so the tempos of the auditory rhythms were chosen to be multiples of this number. The *training stimulus* consisted of a snare drum timbre *downbeat* background presented every 1959 msec at 60 dB(A), a slapstick timbre *microbeat* background presented every 326.5 msec at 50 dB(A) such that 6 microbeats occurred within each period defined by the downbeats, and an ambiguous familiarization rhythm pattern of interest at 60 dB(A) that was superimposed on the background beats. The rhythm pattern of interest consisted of four snare drum sounds with stimulus onset asynchronies (SOAs) of 653–326.5–

326.5–653 msec (where the first sound coincided with the background downbeat) resulting in a rhythm of sound-rest-sound-sound-sound-rest (see Fig. 1). Note that this rhythm pattern is ambiguous because it can be perceived either as consisting of three groups of two beats, with every second beat accented as in a march (SOUND-rest-SOUND-sound-SOUND-rest), or as two groups of three beats, with every third beat accented, as in a waltz (SOUND-rest-sound-SOUND-sound-rest), even though these accents are not physically present. The sounds can be heard at <http://psycserv.mcmaster.ca/ljt/research.htm>.

The two auditory test stimuli were identical to the training rhythm described above, except that the rhythm pattern of interest was disambiguated by physically accenting some sounds relative to others by playing them louder (60 vs. 55 dB). Specifically, for the *duple rhythm*, every second beat was physically accented as in a march (SOUND-rest-SOUND-sound-SOUND-rest), and for the *triple rhythm*, every third beat was physically accented as in a waltz (SOUND-rest-sound-SOUND-sound-rest). In all cases, the beat onset-to-onsets fell within the optimal range for tempo discrimination (Fraisse, 1982; Baruch and Drake, 1997).

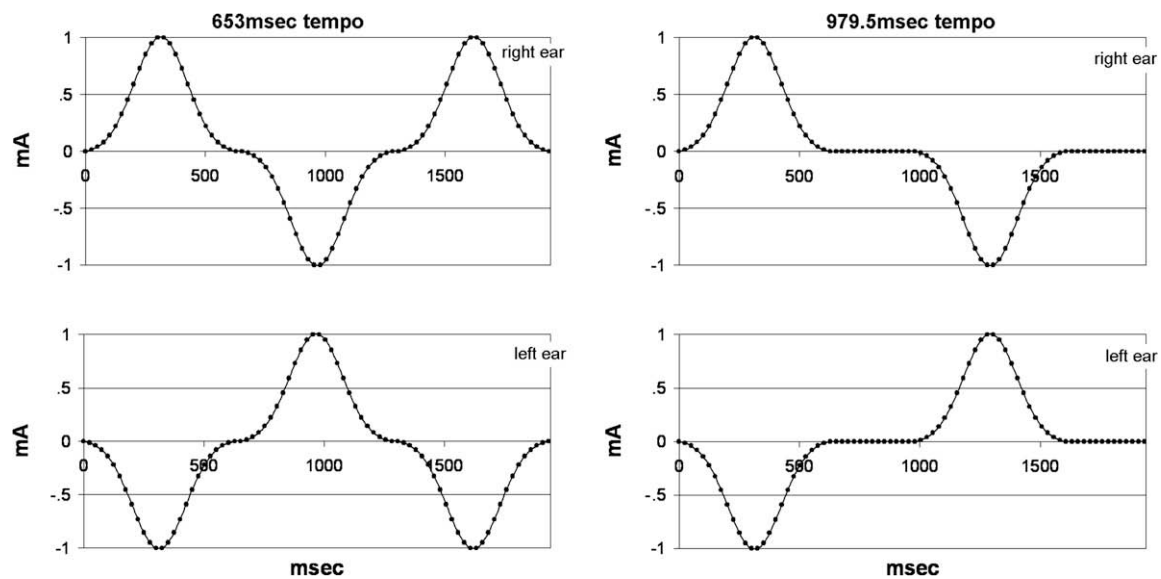
#### 2.2.2. Vestibular stimulus

The vestibular stimulation consisted of a small current applied to electrodes on the mastoid process behind the ears (see Section 2.3). The current waveform consisted of Gaussian-shaped pulses with a peak amplitude of 1 mA and a standard deviation of 100 msec. Positive and negative pulses were presented alternately such that they were out of phase in the two ears (Fig. 2). Three tempos were used for the peak-to-peak time interval between stimulations, 653 msec (corresponding to the duple auditory stimulus rate), 979.5 msec (corresponding to the triple auditory stimulus rate), and 816 msec (midway between the duple and triple tempos, used to calibrate individual phase shifts between vestibular stimulation and perceived head movement, see Section 2.5, below). During pilot testing, we examined different waveforms for the vestibular current stimulation, including impulse waves, triangular waves, and sine waves, and found that the Gaussian waveform produced the clearest sensation of the head moving from side to side.

### 2.3. Apparatus

The vestibular stimulus was generated by a Good Vibrations Engineering Ltd. Galvanic Vestibular Stimulation system (GVS). For the experimental group, an electrode was attached to the mastoid process under each of the participant's ears. For the control group, an electrode was attached to each of the participant's elbows (Fig. 3). The electrode was a 2-inch round Proflex CC carbon conductor electrode (Canadian Medical Products Ltd, F2020PF). This electrode material and shape was selected during pilot testing as providing the most comfort and the best sensation of head movement.

The auditory stimuli were created using Cakewalk with the snare drum (#229) and slapstick (#244) voices on a Roland 64-Voice Synthesizer Module, digitized with Cool Edit 2000 on a personal computer using an AOpen AW-840 4-channel PCI sound card, and presented by a Power Macintosh 7300/180



**Fig. 2 – The galvanic stimulus delivered to the vestibular system. The left panels represent the Gaussian-shaped electrical stimuli at the 653 msec tempo (duple time) and the right panel the 979.5 msec tempo (triple time). For each panel the stimuli are out of phase across the right (top panels) and left (bottom panels) ears.**

computer through a Denon PMA-480R amplifier to two audiological GSI speakers located inside a large Industrial Acoustics Co. sound-attenuating booth with a sound floor of 25 dB(A). The sound booth was set up so that the participant sat between the two speakers, and auditory stimuli were always presented from both speakers.

The experiment was controlled from the GVS system, with software running on an IBM Thinkpad 760ED laptop. At the same time that a vestibular stimulus was sent to the subject, a trigger was sent to the Power Macintosh through a custom-built interface consisting of an opto-isolator. Sounds were presented from the Macintosh using a custom software program that ensured accurate timing.

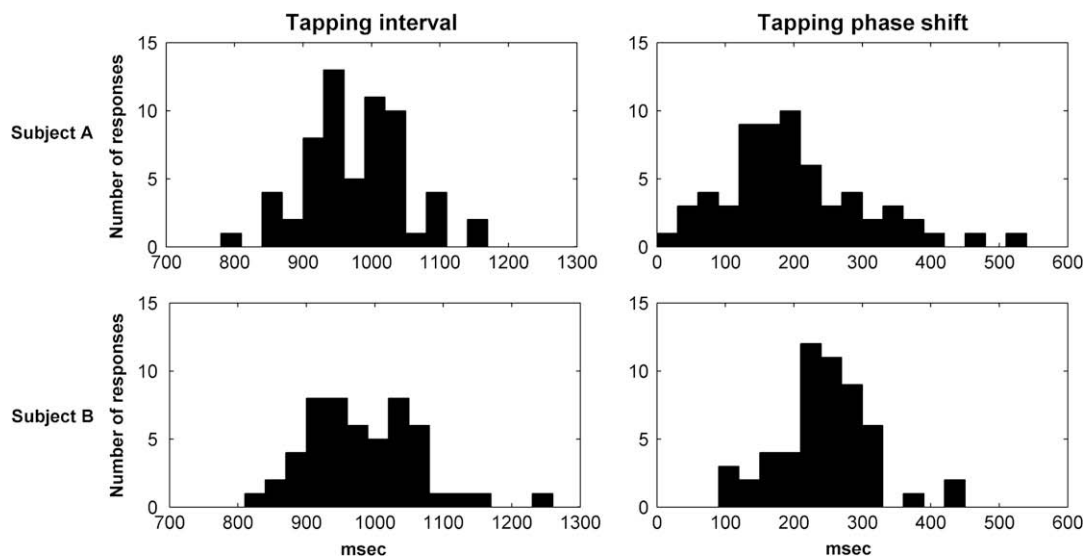
#### 2.4. Pilot test of phase relations between vestibular stimulation and perceived head movement

Pilot testing revealed that with vestibular stimulation of alternating positive and negative Gaussian waveforms that were out of phase between the ears, subjects experienced a side-to-side movement of the head, even though the head actually remained stationary. However, the phase relation between the vestibular stimulation and the time at which the head was perceived to be maximally displaced to the right and to the left varied considerably from person to person. It was essential that the time of perceived maximal head displacement corresponded to the onsets of auditory beats in the main experiment, so it was necessary to characterize the extent of these individual differences and to compensate for them on an individual basis. In order to do this, 5 pilot subjects were given vestibular stimulation alone (i.e., no concurrent sound) in an isochronous rhythm sequence, with the peaks of the vestibular stimulation occurring at 653 msec intervals in one condition (corresponding to the duple auditory stimulus rate)

and at 979.5 msec intervals (corresponding to the triple auditory stimulus rate) in a second condition. Subjects were asked to tap along with their perceived side-to-side head movement on a response pad (EGI 200) at the precise times at which their head was maximally displaced to the right and to the left. Subjects were monitored to ensure that they did not make any overt head movements. Fig. 4 shows histograms of the onset-to-onset times of tapping intervals for two individual subjects for the 979.5 tempo. As can be seen, responses centred on the vestibular inter-stimulus separation times, indicating that the head was perceived to move at the same rate as the vestibular stimulation. Fig. 4 also shows the phase offset between vestibular peak stimulation and the timing of the subjects' taps. As can be seen, this phase varied considerably between subjects. Because we did not want to fatigue the vestibular responses of subjects by running them through a long pilot procedure, a short test was developed to determine the



**Fig. 3 – Vestibular and control conditions. In the vestibular condition, galvanic stimulation was delivered to the vestibular nerve (left panel) and in the control condition it was delivered to the elbows (right panel).**



**Fig. 4 – Distribution of tapping responses synchronized to illusory perceived head movements evoked by galvanic stimulation illustrated by the responses of two representative subjects. Bin widths are 30 msec. Stimuli were delivered using the pattern illustrated in Fig. 2 with a separation of 979.5 msec. The distributions of inter-tap intervals (left panels) were centred close to the vestibular stimulation rate at 978 msec ( $SD = 72$ ) and 983 msec ( $SD = 80$ ) for each subject. Phase relations (delay in msec between the stimulus and tap, right panels) showed variability across subjects, peaking at 200 msec ( $SD = 102$ ) and 245 msec ( $SD = 71$ ) in subjects A and B, respectively.**

approximate phase relation for each individual subject as described below.

## 2.5. Procedure

Subjects were first given a pre-test to determine their individual phase relation between vestibular stimulation and their maximal perceived side-to-side head movement. They were then given a questionnaire about their musical background. This was followed by a familiarization phase in which the ambiguous auditory rhythm pattern was presented concurrently with vestibular stimulation on every second (duple familiarization condition) or on every third (triple familiarization condition) beat of the ambiguous auditory rhythm. In this phase, the control subjects experienced stimulation of their elbows instead of their vestibular nerve. Subjects were monitored to ensure that they did not make any overt head movements. Finally, subjects were given an auditory-alone test with the disambiguated duple and triple auditory stimuli (no vestibular stimulation) and asked to choose which sounded most like what they had heard during the familiarization phase. The entire experiment lasted less than 1 h.

All subjects heard exactly the same auditory stimuli throughout the experiment. If there were multisensory interactions between the vestibular and auditory systems, subjects who experienced duple vestibular stimulation should perceive the ambiguous auditory pattern as a march, and hence choose the duple auditory stimulus as sounding most like what they heard during familiarization. On the other hand, those experiencing triple vestibular stimulation would be expected to perceive the ambiguous

auditory pattern as a waltz, and choose the triple auditory stimulus as sounding most like what they heard during familiarization. Each of these phases is described in detail below.

### 2.5.1. Individual phase shift measurement

The electrodes were applied behind the ears in the experimental group and to the elbows in the control group. Subjects in the experimental group were told that they might feel their head move from side to side while they heard a repeating tone. Subjects in the control group were told that they might feel a tingling in their elbows. Experimental group subjects were to determine whether the tone came before or after the time at which they perceived their head to be maximally displaced to the right and left. Control group subjects were to determine whether the tone came before or after the tingling they felt in their elbows. A 100 msec 60 dB pure tone at 500 Hz was played repeatedly with an SOA of 816 msec (half way between the two possible stimulation rates used during the experiment that followed) and the vestibular/elbow stimulation was applied at the same rate for 30 sec. Initially, the tone was presented 200 msec offset from the peak of the electrical stimulation. Subjects were asked to indicate whether the sound came before or after the point of maximal perceived head displacement/elbow tingling or whether the two stimuli were concurrent. The phase was then adjusted until the sound and perceived head movement/elbow tingling were reported to be concurrent. Five possible phase shift values were used, 150, 170, 200, 230, 260, and 290, covering the range of phase values measured in the pilot study. Phase shifts used ranged between 170 and 260 msec with a mean of 205 msec ( $SD = 33.3$ ).

### 2.5.2. Familiarization

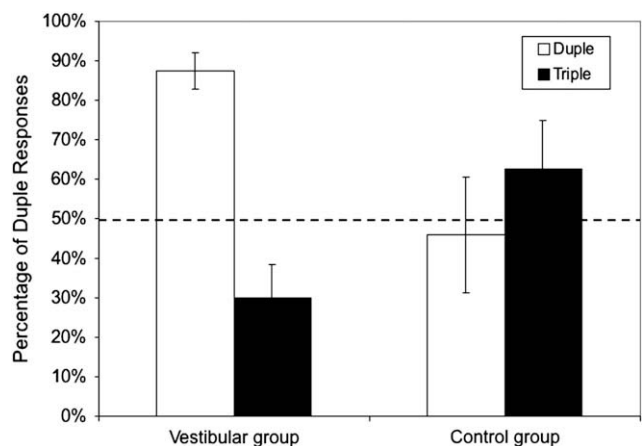
All subjects listened to one minute of the ambiguous auditory familiarization rhythm. Subjects in the experimental condition received concurrent vestibular stimulation while those in the control group received concurrent elbow stimulation. Within each group, half received duple vestibular/elbow stimulation (i.e., with an SOA of 653 msec, individually phase shifted to be perceived as concurrent with every second beat of the auditory pattern) and half received triple vestibular/elbow stimulation (i.e., with an SOA of 979.5 msec, individually phase shifted to be perceived as concurrent with every third beat of the auditory pattern). All subjects were tested with eyes closed. Although it is possible that the vestibular stimulation caused some nystagmus, such eye movements are unlikely to influence perceived auditory rhythm directly. The brain largely monitors such eye movements via an efference copy of the vestibular driving signal rather than feed forward proprioceptive input from eye muscles (Bridgeman and Stark, 1991). If the vestibular stimulation we used were to have an effect on the interpretation of auditory rhythm patterns, it would be even more remarkable if it were mediated by eye movements than if it were mediated directly through the vestibular system. The control condition with the electrodes on the elbows would also likely cause some small muscular activity. A lack of effect in this condition would suggest that small muscle movements are insufficient to influence auditory pattern perception.

### 2.5.3. Testing

Immediately following the training phase, the participant was given eight two-alternative forced-choice trials, with auditory presentation alone. Each trial contained a duple and a triple test stimulus. Both of the duple and triple test stimuli were identical to the auditory familiarization stimulus, except that physical accents (see Section 2.2) were present on every second beat for the duple test stimulus, and on every third beat for the triple. The two stimuli on each trial were presented in random order for one cycle or 1959 sec each, separated by 1000 msec. On the first trial, the order was counterbalanced, so that half of the participants in each condition heard the duple rhythm first, and half heard the triple rhythm first. Participants were instructed to choose which of the two stimuli was the same as, or most similar to, the rhythm they had heard in the training phase, and their verbal responses were recorded by the experimenter.

## 3. Results

After vestibular entrainment with 653 msec intervals (corresponding to the beats of the duple interpretation), 87.5 percent of judgements of the rhythm of the ambiguous auditory stimulus were for the duple interpretation; after entrainment with 979.5 msec intervals (corresponding to the beats of the triple interpretation), 70.0 percent of judgements were for the triple interpretation. An ANOVA, with the independent variables experimental condition (vestibular, elbow) and electrical familiarization stimulation (duple, triple) and the dependent variable the proportion of responses in which the duple



**Fig. 5 – Results. Percentage of times subjects identified the ambiguous auditory pattern as duple for each of the four conditions. Open bars after stimulation at the duple frequency, shaded bars after stimulation at the triple frequency. Left pair of bars after galvanic stimulation of the vestibular system (see Fig. 3, left panel), right pair of bars after control stimulation of the elbows (see Fig. 3, right panel). The horizontal dotted line represents chance responding (50%). Error bars represent the standard error of the mean.**

auditory-alone rhythm was chosen as most similar to what was heard during familiarization, revealed only a significant interaction between experimental condition and familiarization stimulation,  $F(1, 19) = 11.3$ ,  $p < .003$ . As can be seen in Fig. 5, stimulation of the elbows (control condition) at either frequency during familiarization did not affect whether subjects perceived the ambiguous auditory stimulus in duple or triple form in any systematic way,  $p > .4$ . On the other hand, vestibular stimulation during familiarization had a significant effect on whether subjects perceived the ambiguous auditory stimulus in duple (march) or triple (waltz) form,  $t(10) = 6.3$ ,  $p < .0001$ . There were no significant correlations between musical training and performance for either group.

## 4. Discussion

Previous studies have shown that physical movement plays a critical role in entraining and disambiguating a musical rhythm (Phillips-Silver and Trainor, 2005, 2007). The present study shows that this effect is mediated through the vestibular system. Galvanic stimulation of the vestibular system caused, on average, 79% of perceptual judgements to move towards the auditory interpretation entrained by the vestibular stimulation, whereas control stimulation caused no such shift in auditory pattern perception. It is well known that music makes us move, but these studies show that the act of feeling a rhythm is an interactive process: hearing a rhythm evokes physical movement and the resulting vestibular stimulation also influences the auditory interpretation of the rhythm.

The lack of correlation between years of musical training and size of the vestibular influence on audition, in

conjunction with previous findings of movement–auditory interactions in infants (Phillips-Silver and Trainor, 2005), suggests that vestibular influence on auditory processing might arise early in development and not depend on any special experience. The vestibular system is a very primitive system that emerges early in both phylogeny and ontogeny and that determines the organization and development of the other senses. Given the primal role of the vestibular system it is perhaps not surprising that it has such a fundamental influence on auditory perception. Phylogenetically the vestibular system was the first sensory system to develop in evolution (Walls, 1962) and ontogenetically it is the first system to develop in the womb (Romand, 1992), suggesting that a sense of orientation in the gravitational field is more fundamental to perception than is vision and hearing. Indeed vestibular experience appears to be fundamental even for social–emotional development in primates. Vestibular input through movement experience can partially ameliorate the detrimental effects of maternal separation in infant rhesus monkeys raised by cloth-covered surrogate mother cylinders (Mason and Berkson, 1975), as in Harlow's famous experiments (Mason and Harlow, 1958).

An important role for the vestibular system, in addition to its fundamental role in providing orientation information, arises because it is dynamically sensitive in the range of biological rhythms, especially those in the frequency range of locomotion (Wilson and Jones, 1979). Biological rhythms are pervasive and exist at many different temporal scales. Two of the most likely sources of biological rhythm at the time scale of musical rhythm are locomotion and heart beat. The tempo range of locomotion movements corresponds closely to the range over which a beat can be felt – pulse separations of between about 300 and 900 msec (e.g., Fraisse, 1982; Clarke, 1999) and heart rate varies from about 60–150 beats/min, corresponding to a very similar range of 400–1000 msec. Both of these rhythms can be directly sensed by the vestibular system. The large physical motions associated with walking and running of course excite the vestibular system, but even the tiny movements of the head caused by the pulse are also detected (Crawford, 1952). Developmentally, both of these biological rhythms are experienced by the fetus as the mother locomotes through the environment. But at what stage of development, and where in the nervous system does vestibular input connect with auditory rhythmic experience?

Although the auditory and vestibular end organs are anatomically close and functionally similar, sounds must be at least 95 dB before they directly stimulate the vestibular system (Todd and Cody, 2000). As the sounds in this experiment are much quieter than this, the auditory–vestibular interaction that we have demonstrated must occur more centrally. There is a potential concern that our galvanic stimulation might have had a direct effect on the auditory portion of the VIII nerve (Bucher et al., 1998). However, this is very unlikely because the currents used were very small and none of our subjects reported hearing any sounds. The relevant auditory–vestibular interactions are more likely to occur more centrally.

Recent evidence indicates that auditory and vestibular information may interact at many levels of the nervous system, including the dorsal cochlear nuclei (DCN) (Oertel

and Young, 2004), and the posterior parietal cortex (PPC) (Bremmer, 2005; Colby et al., 1993; Lewald et al., 2002; Lewis and Van Essen, 2000; Schlack et al., 2005, 2002). Models of rhythmic movement have largely focused on interactions between PPC, cerebellum and prefrontal cortex (e.g., Todd et al., 2002). The cerebellum in particular has been implicated in the processing of auditory rhythm (Penhune et al., 1998; Parsons, 2003; Griffiths, 2003; but see Molinari et al., 2005). The cerebellum is a major recipient of vestibular input (e.g., Suzuki and Keller, 1982). The role of the cerebellum in timing has been conceptualized not as a clock or counter but simply as the structure that provides the necessary circuitry for the sensory systems to extract temporal information and for the motor system to learn to produce a precisely timed response (Penhune et al., 1998). The effect of vestibular stimulation of auditory rhythm, then, is likely to provide a similar enhancement to the auditory sensory signal on selected strong beats as would an increase in auditory intensity.

The ventral intraparietal area (VIP) of the PPC is of particular interest as well because single cell recordings in the Macaque monkey show that it responds to both auditory (Schlack et al., 2005) and vestibular (Colby et al., 1993; Bremmer et al., 2001; Schlack et al., 2002) input as well as to visual and somatosensory stimulation. 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, offers an account of the auditory–vestibular connections that may underlie our findings of multisensory interactions between movement and the perception of auditory rhythm (see Phillips-Silver and Trainor, 2008).

The locus of vestibular–auditory interactions could be determined in a number of ways. fMRI studies comparing strengths of activation across different regions for auditory alone, vestibular alone, and combined auditory–vestibular stimulation would be informative. Studies of whether patients with lesions in different auditory–vestibular convergence sites experience the influence of vestibular stimulation on auditory rhythm processing would also give information about which brain regions are critical. Finally, if vestibular–auditory effects are cortically mediated, no multisensory effect would be expected in infants two months of age and younger, as the auditory cortex is not mature enough at this stage to support complex processing (Moore and Guan, 2002).

---

## 5. Conclusions

Musical rhythm patterns elicit physical movement from head bobbing and foot tapping to all-out dancing. Our previous work showed that movement of the body can influence the auditory perception of the metrical structure of rhythm. The present paper demonstrates that this effect is mediated by vestibular stimulation and can be recreated in the absence of physical movement by artificially stimulating the vestibular nerve. Most likely, several regions of vestibular–auditory convergence are involved in some way in rhythm processing and it will be for future research to discover the specific role of each region.



## Acknowledgements

This research was supported by grants from NSERC (Canada) to Laurel Trainor and Laurence Harris. We thank Jessica Phillips-Silver for discussions of the stimuli and Lisa Hotson for organizational assistance.

## REFERENCES

- Arom S. *African Polyphony and Polyrythm*. Cambridge: Cambridge University Press, 1991.
- Baruch C and Drake C. Tempo discrimination in infants. *Infant Behaviour and Development*, 20: 573–577, 1997.
- Bremmer F. Navigation in space- the role of the macaque ventral intraparietal area. *The Journal of Physiology*, 566: 29–35, 2005.
- Bremmer F, Schlack A, Duhamel J-R, Graf W, and Fink GR. Space coding in primate posterior cortex. *Neuroimage*, 14: S46–S51, 2001.
- Bridgeman B and Stark L. Ocular proprioception and efference copy in registering visual direction. *Vision Research*, 31: 1903–1913, 1991.
- Bucher SF, Dieterich M, Wiesmann M, Weiss A, Zink R, Yousry TA, et al. Cerebral functional magnetic-resonance-imaging of vestibular, auditory, and nociceptive areas during galvanic stimulation. *Annals of Neurology*, 44: 120–125, 1998.
- Buys E. Beitrag zum studium des galvanischen nystagmus mit hilfe der nystagmographie. *Mschr Ohrenheilk*, 43: 801–803, 1909.
- Clarke EF. Rhythm and timing in music. In Deutsch D (Ed), *The Psychology of Music*. 2nd ed. San Diego: Academic Press, 1999: 473–500.
- Colby CL, Duhamel JR, and Goldberg ME. Ventral intraparietal area of the macaque: anatomic location and visual response properties. *Journal of Neurophysiology*, 79: 126–136, 1993.
- Crawford J. Living without a balancing mechanism. *The New England Journal of Medicine*, 246: 458–460, 1952.
- Cross I. Music, cognition, culture and evolution. In Peretz I, and Zatorre R (Eds), *The Cognitive Neuroscience of Music*. New York: Oxford, 2003: 42–56.
- Cutieta RA. *Raising Musical Kids: A Guide for Parents*. New York: Oxford University Press, 2001.
- Drake C, Penel A, and Bigand E. Tapping in time with mechanically and expressively performed music. *Music Perception*, 18: 1–23, 2000.
- Fraisse P. Rhythm and tempo. In Deutsch D (Ed), *The Psychology of Music*. New York: Academic Press, 1982: 149–180.
- Goldberg JM, Smith CE, and Fernandez C. Relation between discharge regularity and responses to externally applied galvanic currents in vestibular nerve afferents of the squirrel monkey. *The Journal of Neurophysiology*, 51: 1236–1256, 1984.
- Griffiths TD. The neural processing of complex sounds: Musical brain substrates. In Peretz I, and Zatorre RJ (Eds), *The Cognitive Neuroscience of Music*. New York, USA: Oxford University Press, 2003: 168–177.
- Jaques-Dalcroze E. *Method of Eurhythmics: Rhythmic Movement*, vol. 1. London: Novello and Co, 1920.
- Juntunen ML and Hyvönen L. Embodiment in musical knowing: how body movement facilitates learning with Dalcroze Eurhythmics. *British Journal of Music Education*, 21: 199–214, 2004.
- Lerdahl F and Jackendoff R. *A Generative Theory of Tonal Music*. Cambridge: MIT Press, 1983.
- Lewald J, Foltys H, and Töpper R. Role of the posterior parietal cortex in spatial hearing. *The Journal of Neuroscience*, 22: 1–5. RC207, 2002.
- Lewis J and Van Essen DC. Corticocortical connections of visual, sensorimotor and multimodal processing areas in the parietal lobe of the macaque monkey. *Journal of Computational Neurology*, 428: 112–137, 2000.
- Mars F, Vercher JL, and Popov K. Dissociation between subjective vertical and subjective body orientation elicited by galvanic vestibular stimulation. *Brain Research Bulletin*, 65: 77–86, 2005.
- Mason WA and Berkson G. Effects of maternal mobility on the development of rocking and other behaviors in rhesus monkeys: a study with artificial mothers. *Developmental Psychobiology*, 8: 197–211, 1975.
- Mason WA and Harlow HF. Formation of conditioned responses in infant monkeys. *Journal of Comparative and Physiological Psychology*, 51: 68–70, 1958.
- Molinari M, Leggio MG, De Martin M, Cerasa A, and Thaut M. Neurobiology of rhythmic motor entrainment. *Annals of the New York Academy of Science*, 999: 313–321, 2003.
- Molinari M, Leggio MG, Filippini V, Gioia MC, Cerasa A, and Thaut MH. Sensorimotor transduction of time information is preserved in subjects with cerebellar damage. *Brain Research Bulletin*, 67: 448–458, 2005.
- Moore JK and Guan YL. Cytoarchitectural and axonal maturation in human auditory cortex. *Journal of the Association for Research in Otolaryngology*, 2: 297–311, 2002.
- Oertel D and Young ED. What's a cerebellar circuit doing in the auditory system? *Trends in Neuroscience*, 27: 104–110, 2004.
- Parsons LM. Exploring the functional neuroanatomy of music performance, perception, and comprehension. In Peretz I, and Zatorre RJ (Eds), *The Cognitive Neuroscience of Music*. New York, USA: Oxford University Press, 2003: 247–268.
- Patel AD, Iversen JR, Bregman MR, Schultz I, and Schulz C. Investigating the human-specificity of synchronization to music. Presented at the International Conference on Music Perception and Cognition, Sapporo, Japan, August 2008.
- Penhune VB, Zatorre RJ, and Evans AC. Cerebellar contributions to motor timing: a PET study of auditory and visual rhythm reproduction. *Journal of Cognitive Neuroscience*, 10: 752–765, 1998.
- Phillips-Silver J and Trainor LJ. Feeling the beat: movement influences infant rhythm perception. *Science*, 308: 1430, 2005.
- Phillips-Silver J and Trainor LJ. Vestibular influence on auditory metrical interpretation. *Brain and Cognition*, 67: 94–102, 2008.
- Phillips-Silver J and Trainor LJ. Hearing what the body feels: auditory encoding of rhythmic movement. *Cognition*, 105: 533–546, 2007.
- Repp BH. Sensorimotor synchronization: a review of the tapping literature. *Psychonomic Bulletin and Review*, 12: 969–992, 2005.
- Romand R (Ed), *Development of Auditory and Vestibular Systems*. New York: Elsevier, 1992.
- Schachner A, Brady TF, Pepperberg IM, and Hauser MD. Entrainment to auditory rhythms in vocal learning bird species. Presented at Neurosciences and Music III, Montreal, Canada, 2008.
- Schlack A, Hoffman KP, and Bremmer F. Interaction of linear vestibular and visual stimulation in the macaque ventral intraparietal area (VIP). *European Journal of Neuroscience*, 16: 1877–1886, 2002.
- Schlack A, Sterbing-D'angelo SJ, Hartung K, Hoffmann K-P, and Bremmer F. Multisensory space representations in the macaque ventral intraparietal area. *The Journal of Neuroscience*, 25: 4616–4625, 2005.
- Suzuki DA and Keller EL. Vestibular signals in the posterior vermis of the alert monkey cerebellum. *Experimental Brain Research*, 47: 145–147, 1982.
- Todd NP, Cousins R, and Lee CS. The contribution of anthropometric factors to individual differences in the perception of rhythm. *Empirical Musicology Review*, 2, 2007.
- Todd NP, Lee CS, and O'Boyle DJ. A sensorimotor theory of temporal tracking and beat induction. *Psychological Research*, 66: 26–39, 2002.

- 
- Todd NPM. The kinematics of musical expression. *Journal of the Acoustical Society of America*, 97: 1940–1949, 1995.
- Todd NPM and Cody FW. Vestibular responses to loud dance music: a physiological basis of the “rock and roll threshold”? *Journal of the Acoustical Society of America*, 107: 496–500, 2000.
- Trainor LJ. Do preferred beat rate and entrainment to the beat have a common origin in movement? *Empirical Musicology Review*, 2: 17–20, 2007.
- Wallin NL, Merker B, and Brown S. *The Origins of Music*. Cambridge, MA: MIT Press, 2000.
- Walls GL. The evolutionary history of eye movements. *Vision Research*, 2: 69–79, 1962.
- Wardman DL, Taylor JL, and Fitzpatrick RC. Effects of galvanic vestibular stimulation on human posture and perception while standing. *The Journal of Physiology*, 551: 1033–1042, 2003.
- Wilson VJ and Jones GM. *Mammalian Vestibular Physiology*. New York: Plenum, 1979.
- Zink R, Stedden S, Weiss A, Brandt T, and Dieterich M. Galvanic vestibular stimulation in humans – effects on otolith function in roll. *Neuroscience Letters*, 232: 171–174, 1997.