TOPICS
TOPICS IN COGNITIVE SCIENCE

Topics in Cognitive Science 4 (2012) 653–667

ISSN: 1756-8757 print / 1756-8765 online
DOI: 10.1111/j.1756-8765.2012.01215.x

Music Perception and Cognition: A Review of Recent Cross-Cultural Research

Catherine J. Stevens

MARCS Institute and School of Social Sciences and Psychology, University of Western Sydney

Received 15 September 2010; received in revised form 15 May 2011; accepted 16 March 2012

Abstract

Experimental investigations of cross-cultural music perception and cognition reported during the past decade are described. As globalization and Western music homogenize the world musical environment, it is imperative that diverse music and musical contexts are documented. Processes of music perception include grouping and segmentation, statistical learning and sensitivity to tonal and temporal hierarchies, and the development of tonal and temporal expectations. The interplay of auditory, visual, and motor modalities is discussed in light of synchronization and the way music moves via emotional response. Further research is needed to test deep-rooted psychological assumptions about music cognition with diverse materials and groups in dynamic contexts. Although empirical musicology provides keystones to unlock musical structures and organization, the psychological reality of those theorized structures for listeners and performers, and the broader implications for theories of music perception and cognition, awaits investigation.

Keywords: Empirical cognitive ethnomusicology; Entrainment; Expectations; Meter; Multimodal perception; Pitch; Rhythm; Universals

1. Preamble

World musics provide a laboratory for the investigation of music perception and cognition where there has been little or no prior exposure to a particular musical system. Such cross-cultural research is important for at least three reasons. First, as the West encroaches, there is an irreversible loss of cultural diversity (Huron, 2008) and with it

Correspondence should be sent to Catherine J. Stevens, MARCS Institute and School of Social Sciences and Psychology, University of Western Sydney, Locked Bag 1797, Penrith, NSW 2751, Australia. E-mail: kj.stevens@uws.edu.au

knowledge of different musical systems and psychological processes. Second, because theories of perception and cognition have been conceived in Western settings and within a monoculture, psychological theories need to be evaluated and challenged in a variety of contexts. Third, the bias to the West has inhibited the takeup of theory and empirical findings that arise from the way music is created, performed, perceived, used, and learned in diverse cultural settings. An extreme example of this has been a concentration during the 20th century on Western tonal art music that is performed for a seemingly passive audience or listened to via audio recording. Such a narrow view would not have occurred if we had considered situations where music, movement, dance, and visual art are inseparable and music is social, dynamical, and interpersonal (see Barwick, 1989; Blacking, 1995; Hanna, 1979; Nettl, 2005; Titon, 1992). Here, we report cross-cultural studies of music perception and cognition with an emphasis on research conducted within the last decade.

1.1. Music and movement: Diversity and processes; universals and features

Some documentation of world musics has led to the collation of candidate universals in music (e.g., Brown & Jordania, in press; Carterette & Kendall, 1999). Universals imply static features of static environments. Not only does the search for universals and similarities overlook informative differences, there is also a logical problem in documenting definitive generalities rather than particular instances. One alternative is to document and analyze diversity. Evans and Levinson (2009) promote such an approach in linguistics—surveying linguistic diversity to show how differently languages can be structured at phonetic, phonological, morphological, syntactic, and semantic levels. As diverse musical cultures are endangered by globalization (Huron, 2008), now is the time for detailed survey of differences in musical structures with some ways to start provided by Clayton, Herbert, and Middleton (2003) and Nettl (2005). A complementary approach is to recognize musical *processing* over features and content—that is, the psychological processes implicated in engaging in musical behavior rather than the content of musical knowledge (Harwood, 1976; Meyer, 1960).

Concentration on the musical stimulus makes sense when we consider that the way humans construct a musical stimulus or environment is likely to reflect perceptual and cognitive constraints (McAdams, 1989). However, even in the Gibsonian ecological psychology sense, it is essential to consider the way a human interacts with the environment and what environmental structures, objects, and events *afford*. The psychological relationship of cognition comprises the person (his or her acquired knowledge, beliefs, and thoughts) and the stimulus/environment (see Petocz, Keller, & Stevens, 2008). Both elements of the relation may be scrutinized, but it is the interaction that is informative. Studying, for example, features of a musical stimulus that give rise to emotional response neglects the active role of the perceiver and the knowledge that the person brings to the situation. From this perspective, cues to emotional response are not *in* the music but a product of the dynamic interaction between person and stimulus. In the following overview, we will see that there are some but not many studies that examine auditory–visual–kinesthetic processes in perceiving and making music and that recognize the active role of the human perceiver.

2. Effects of enculturation on music perception and cognition

There is a steady increase in the studies of the way listeners perceive the music of other cultures and, in particular, the way often through mere exposure, listeners have been enculturated or become sensitive to grouping, segmentation, and tonal and temporal regularities of their musical environment.

2.1. Grouping and segmentation

A strand of research triggered by Bregman's pioneering account of auditory scene analysis (Bregman, 1990) was the search for principles of perceptual organization in the auditory domain. Using a cross-cultural paradigm, Iversen, Patel, and Ohgushi (2008) demonstrated that perceived auditory grouping is associated with experience. When native English and Japanese speakers listened to simple rhythmic sequences of tones, the most common Japanese preference was for long–short grouping, a grouping rarely selected by English speakers. The dominant language of each cultural group is used as an explanation; specifically, statistical learning of the duration patterns of common rhythmic units in speech shaping low-level grouping biases. The authors hypothesize prevalence of short–long pattern in English and long–short in Japanese with respect to function–content word order. Further research is needed to scrutinize whether language experience *causes* perceptual grouping.

In a series of experiments using event-related potential (ERP) and functional magnetic resonance imaging techniques, Nan et al. have investigated the universality of phrase perception in music and the extent to which it is modified by cultural differences. For example, Nan, Knösche, and Friederici (2006) asked Chinese and German musicians to perform a categorization task using Chinese and Western music. Chinese music was familiar to only the Chinese participants, whereas Western music was familiar to both groups. Pauses were manipulated between the two phrases of short melodies; the perception of phrase boundaries occurred for both groups in response to both kinds of music. The authors liken the observed ''closure positive shift'' (CPS) to a similar effect in language, noting that the most important cues for perception of phrase boundaries are the insertion of a phrase-end pause and the lengthening of the phrase-final syllable or note before the pause. According to Nan et al., the electroencephalography correlate of the CPS is a parietally distributed positive deflection (P2) for phrased versus unphrased items; the ERP component occurs between 450 and 600 ms after the onset of the pause. Differences in N1 between phrased and unphrased items were suggested but not reliable. The results are interpreted as indicative of the perception of phrase boundaries in music as universal.

2.2. Sensitivity to tonal and melodic hierarchies

Music is often generative, multilayered, and hierarchical (Lerdahl & Jackendoff, 1983; Thompson & Schellenberg, 2006; Titon, 1992). Hierarchical structuring of streams of sound, within and across dimensions, is thought to confer a processing advantage (Bharucha, Curtis, & Paroo, 2006). Tonal hierarchies organize tones within a key into stable and unstable

pitches; that is, some tones are perceived as more stable and important than others. Exceptions to hierarchical organization in music include minimalist music (e.g., Philip Glass, Steve Reich) and the cyclical structures of Balinese gamelan (Tenzer, 2006). These cultural variations imply that hierarchy is sufficient but not necessary for music perception and cognition.

In a cross-cultural experiment, Castellano, Bharucha, and Krumhansl (1984) demonstrated that both Indian and American listeners are sensitive to the hierarchical ordering of tones in Indian melodic forms called *rāgs*. Sensitivity to tonal hierarchies has also been established in response to Balinese (Kessler, Hansen, & Shepard, 1984), Scandinavian North Sami yoiks (Eerola, Louhivuori, & Lebaka, 2009; Krumhansl et al., 2000), and Korean Court music (Nam, 1998). These experiments have sampled participants from within and outside those musical cultures.

Hierarchical structures are evident in the melodic dimension of the *ngintaka* song series performed by Antakarinya, Yankunytjatjara, and Pitjantjatjara people of Central Australia (Barwick, 1989). In *ngintaka* songlines, there are three related forms of the basic melodic contour: a linear form, cyclical form, and a transposing cyclical form. The linear melodic form consists of an introductory section, preliminary descent, main descent, and concluding section. Although musicological analyses have brought these organizing principles to light, perception-oriented research awaits with the goal of investigating the perceptual and cognitive reality of such structures. One possible hypothesis is that hierarchical structure is cognitively valid in musical cultures that maintain an oral tradition.

2.3. Tonal, melodic, and stylistic expectations

A piece of music in a familiar genre generates expectations based on knowledge about common features of the genre. In Western tonal music, musical expectations are shaped by rhythmic or metric structure (Large & Jones, 1999), tonal and harmonic structure (Schmuckler, 1989), and melodic structure (Cuddy & Lunney, 1995). Although it is unclear just how broadly these principles apply cross-culturally, it is known from analyses of classical and folk music from different cultures that, statistically, small intervals predominantly descend and large intervals ascend (Vos & Troost, 1989). In support of these statistical regularities, Huron (2006) reports a tendency for large intervals to also ascend in Chinese folk songs, traditional Korean music, Australian Aboriginal music, Ojibway, Pondo, Venda, and Zulu songs. Huron notes that the majority of melodic movements are descending small intervals and he relates this regularity to a similar pattern in spoken language, namely declination. The basis of declination is associated with the exhaustion of air supply in the lungs over the course of a spoken (or sung) utterance.

Employing a recognition-memory paradigm, Curtis and Bharucha (2009) investigated the strength of schematic musical knowledge of Western listeners presented with Western major (culturally familiar) or Indian *thaat* Bhairav (culturally unfamiliar) sets of isochronous tones. Tone sets provided an encoding context when a musically related tone was missing. Participants judged whether a subsequently presented test tone had been presented in the previous set. It was hypothesized that if participants are familiar with the musical culture

underpinning the tone set, then they would tend to make false alarms when the test tone was musically related to the tone set; that is, they would think that the tone had been in the encoding context when it had not. By contrast, participants unfamiliar with the musical culture that underpinned the tone set would not make false alarms to the test tone. Results supported the hypothesis suggesting that internalized cultural knowledge influences musical expectancies in the context of an unfamiliar musical scale system.

Although obtained using a recognition memory task, Curtis and Bharucha (2009) link the effect with differential emotional responses for native and nonnative listeners. Similarly, Huron (2006) posits that implicit, statistical expectations are linked directly to emotional response, eliciting basic neurobiological responses. He argues that two processes, imagination and tension, are related to the anticipation of an event. Three expectation processes—prediction, reaction, and appraisal—are theorized to be related to the outcome of events. The development of musical expectations is theorized to be achieved by statistical learning (e.g., McMullen & Saffran, 2004).

2.4. Temporal hierarchies: Examples and exceptions

Event or temporal hierarchies extend upward from the smallest subdivisions of a beat, to beat level, to a phrase, to large-order forms (Clayton, Sager, & Will, 2005). Meter in the great majority of Western music is regular with beat levels containing even divisions of time and time intervals related by simple ratios 1:1, 2:1. Music from Balkan, Asian, and African regions reflect meters that are irregular. Most Western adult listeners have difficulty in accurately producing and perceiving such meters. For example, stimuli in which the longer interval is less than double the duration of the shorter interval tend to be reproduced with a long/short ratio of 2:1, that is, simplified (Fraisse 1946 cited in Povel, 1981). It appears, then, that enculturation also occurs in response to temporal structures. Soley and Hannon (2010) demonstrated early preferences for meters of one's culture. American infants preferred Western over Balkan meter, whereas Turkish infants familiar with Western and Balkan meter showed no preference. In subsequent experiments, both groups of infants preferred Western and Balkan meter to an arbitrary complex ratio meter. Soley and Hannon conclude that early preferences are influenced by both culture-specific experience and a general preference for simplicity.

Just as there are cultural exceptions to hierarchical organization of tones (see Section 2.2), the pulsation of some African musical styles differs from the hierarchical concept of measure in Western classical music and the related concepts of meter, strong, and weak beats (Blacking, 1973; Carterette & Kendall, 1999; Iyer, 1998, 2002; Magill & Pressing, 1997). Pulsations—a sequence of isochronous temporal units—can be realized as a beat that is the analog of the tactus of Western tonal music. In west African music, for example, the main beat and its metric grouping are articulated with ''suggestion and complexity'' rather than accentual reinforcement (Iyer, 1998). The meter is encoded in the rhythm itself; it is unambiguous but culturally specific. Iyer (1998) gives the example of a standard bell pattern that would be heard phrased in three different ways by three different cultural groups.

Examining *tal*, a concept in North Indian (Hindustani) music broadly equivalent to meter, Clayton (1997) notes both similarities and differences. *Tal* is a temporal framework acting as a background for rhythmic design; it is a periodic and hierarchic temporal framework involving the interaction of two or more streams of pulsation. Rhythm in Indian music is interpreted with respect to *tal* with musically untrained Indian listeners inferring *tal* much as musically untrained Western listeners successfully abstract meter.

Musicological accounts of diverse music have helped refine assumptions about music cognition. For example, accelerating rhythms are found in Japanese Gagaku and Tibetan monastic music (Huron, 2006). Accelerating rhythms are not periodic yet are able to be produced and perceived. Huron notes no known instance of accelerating rhythms in Western music before the 20th century. Thus, traditions of the east indicate that periodicity is not necessary for music production/perception.

2.5. Temporal expectations, synchronization, and entrainment

Temporal entrainment is central to music perception and production within and across individuals and groups (e.g., Bispham, 2006; Drake & Bertrand, 2001; Merker, 2000; Merker, Madison, & Eckerdal, 2009; Nozaradan, Peretz, Missal, & Mouraux, 2011; Phillips-Silver, Aktipis, & Bryant, 2010; Phillips-Silver & Keller, 2012; Zentner & Eerola, 2010). Entrainment is a process where two rhythmic processes interact with each other, eventually locking in to a common phase and/or periodicity. A possible mechanism is neural oscillators entraining to external events such as meters or rhythms (Barnes & Jones, 2000; Drake, Jones, & Baruch, 2000; Jones, 2004). Further research is needed to investigate these assumptions of periodicity in the context of asymmetric meters and aperiodic patterns (e.g., Velasco & Large, 2011).

Synchronization to a beat is of current research interest in nonhuman species, in infants, and in and between adults. For example, Patel, Iversen, Bregman, and Schulz (2009) reported synchronization to a beat and adjustment to varying tempo by a sulphur-crested cockatoo (*Cacatua galerita eleonora*). The results conform with a vocal learning and rhythmic synchronization hypothesis (Merker, 2005; Patel, 2006) where entrainment to a musical beat relies on neural circuitry for vocal learning and coupling of auditory and motor circuits. The broad hypothesis is that vocal learning species should be capable of synchronization with a beat.

A sensitive period for humans acquiring rhythmic expectations has been proposed by Hannon and Trehub (2005). The conclusion is based on a developmental study that used a Balkan folk tune stimulus containing 2:1 and 3:2 ratio intervals. Research is emerging that investigates the effect of cultural background on: joint drumming comparing Brazilian and German preschool children (Kirschner, Ilari, & Tomasello, 2010); and the bodily expression of beat and meter in dance movements to South African songs (Himberg & Thompson, 2011). African and European melodies, to which African and European listeners tapped along, were the experimental materials used by Toiviainen and Eerola (2003). No differences in tapping behavior of the two groups of listeners were found in response to European melodies. However, the distribution of tap times and tapping consensus differed in response

to African melodies. European listeners had greater difficulty in synchronizing with African melodies that contained rhythmic motives similar to clave rhythm. A clave rhythm is a two-bar pattern with two accentuations in one bar and three in the other (Finlay, 2011).

The foregoing review points to effects of enculturation similar to perceptual reorganization observed in infants and children in the context of speech perception and production (see Burnham & Mattock, 2010). More critically, we might ask, what has been learned about other cultures and concepts in music? There is scope to broaden or validate theories of cognition based on the form, role, and importance of music in diverse cultures. One development is to consider connections *between* tonal and temporal structures.

2.6. Interconnected tonal and temporal expectations

There has been a tendency for the dimensions of pitch and time, melody and rhythm, to be treated separately in studies of Western music perception and cognition. Considering music perception and production as multimodal and intermodal, embodied and kinesthetic, encourages new ways of thinking about musical dimensions. Turpin (2007a) describes *Akwelye* songs owned by the people of *Arnerre*, a Katetye-speaking group of Central Australia. All songs within a song series have the same melodic structure, and similar rhythmic patterns of text can be sung with different melodic contours. However, substitution of a different text while maintaining the same melody and rhythm would be regarded as incorrect or a different song (Turpin, 2007b). The rhythmic text cycles until the end of the pitch contour is reached. Turpin (2007b) posits that rhythm and melody are separable and independent in Central Australian songs in a way that English songs are not. The cognitive representation of text, rhythm, and melody for the performers of *Akwelye* songs is an open question.

Melody–rhythm interactions are evident in a study of the perception of Arabic improvised music or *taqsīm* (Ayari & McAdams, 2003). Differences in identification of musical elements and segmentation were found between listeners of European and Arabic cultural origins. Both groups segmented music on the basis of salient surface features; Arab listeners also made segmentations defined by subtle modal changes. Melodic reductions revealed that Arabic modes involve not only a tuning system but melodico-rhythmic configurations emblematic of the *maqām* mode.

Given processes of enculturation and acquired perceptual habits, it is challenging to compare diverse music. How do we measure similarity or complexity without the bias of a particular cultural slant? Eerola, Himberg, Toiviainen, and Louhivuori (2006) set out to investigate judged complexity of African and Western folk melodies using intact musical materials incorporating interactions between temporal, tonal, timbral, and intensity dimensions. Both the African and Finnish participant groups were familiar with Western materials and, in response to Western music, their ratings of complexity were similar with respect to musical variables such as entropy of interval distribution, note density, average interval size, accent, coherence, and tonal ambiguity. African melodies were familiar in style to African listeners but not Finnish listeners. Accordingly, there was a lower predictive value of musical variables and group differences with respect to rhythm and structural variables. African participants gave higher complexity ratings to African music and rhythm variables

diverged, suggesting that African melodies contain different types of rhythmic features. The pitch variable category appeared irrelevant for complexity ratings of African melodies by Western listeners. The use of intact materials and the approach adopted by Eerola et al. offers a new way to scrutinize musical difference and universality.

A recurring theme in the present overview of enculturation effects has been the development of perceptual habits from exposure to a particular musical environment. Another trend in current research is to investigate similarities and differences in the processing of music and another universal and culturally varied human behavior—language perception and production.

3. Cultural influences on music and language perception and production

Physical limits imposed by breathing were noted earlier as influences on phrasing (Section 2.1) in music and speech, and declination thought to be associated with a predominance of descending small intervals in melodies (Section 2.3). With regard to rhythm, an association between the prosody of a composer's native language and the structure of composed music has been demonstrated (Patel & Daniele, 2003). A quantitative measure of speech rhythm—the normalized Pairwise Variability Index (nPVI)—was applied to music of 16 composers from France and England. The nPVI is a measure of the patterning of vowel duration in speech. It returns a measure of durational variability calculated from the variation between successive intervals. Patel and Daniele applied the nPVI to durations of successive tones in musical sequences with the obtained values, reflecting durational differences between each pair of tones. French is a syllable-timed language and English a stress-timed language and, as anticipated, English and French instrumental classical music have significantly different nPVI or temporal patterning values in line with differing linguistic durational patterning. The nPVI offers a metric for the analysis of other musics, languages, and their interrelationships.

The Black Atlantic rhythmic diaspora realized in jazz, blues, rock, and hip-hop adopts, according to Pressing (2002), structures shared with speech, especially prosody, conversation interaction, and narrative. Pressing describes devices, including syncopation, displacement, polyrhythm and polymeter, call and response, and speech-based rhythms. His thesis is that the devices establish or enhance a message via increased engagement and focused attention. Stobart and Cross (2000) identified influences of prosody of the Quechua language on Easter songs of Northern Potosí, Bolivia, with the implication that rhythmic and metric structures are culture-specific.

The effect of language background, especially tonal language background, on absolute pitch (AP)—the ability to identify or produce isolated tones in the absence of contextual cues—has attracted recent interest. It is noteworthy that 70% of the world's languages are tonal; that is, differences in pitch serve to distinguish the core meaning of a word (Yip, 2002). Therefore, tonal languages constitute a rich and diverse medium for research. Deutsch, Henthorn, and Dolson (2004) reported that native speakers of tone languages—Mandarin and Vietnamese—showed a precise form of AP in enunciating words.

By contrast, Schellenberg and Trehub (2008) tested Canadian children of Chinese or European background and found that knowledge of a tone language did not affect performance on a task involving long-term memory for pitch level. Schellenberg and Trehub document various explanations for the oft-cited Asian advantage for AP and consider possible genetic and/or cultural differences. The effect of tonal language background on response to pitch in spoken and musical items is a burgeoning area of study (e.g., Bent, Bradlow, & Wright, 2006; Chandrasekaran, Krishnan, & Gandour, 2009; Pfordresher & Brown, 2009; Stevens, Keller, & Tyler, in press).

In many of the studies described so far, the participant groups are ''bimusical.'' Wong, Roy, and Margulis (2009) define bimusicalism as the dual enculturation of memory and emotion. In the Wong et al. study, listeners with significant exposure to Indian and Western traditions were compared with listeners familiar with only Indian or Western culture. Participants completed recognition memory and tension judgment tasks in response to unfamiliar Indian and Western music. Indian and Western listeners showed an in-culture bias, whereas those familiar with both Indian and Western cultures showed equal response to music from both cultures. Future research will likely investigate the degree to which bimusicalism mirrors processes of bilingualism and the explanatory power of the former concept.

4. Emotional response to music

A prolegomenon for cross-cultural studies of emotion and music can be gleaned from Juslin and Västfjäll's (2008) theory of six mechanisms in emotional response to music: brain stem reflex, evaluative conditioning, emotional contagion, visual imagery, episodic memory, and musical expectancy. Evaluative conditioning acknowledges a relational view of emotion and listener perception where emotion is not *in* the music but a relation in which music and listener stand to one another. Conditioning highlights the associations that may be acquired, for instance, between a particular pattern of sound, such as the falling minor third interval (Curtis & Bharucha, 2010) and common associations with and responses to it. Musical expectancy appears to play a role not only as one aspect of musical memory but also in emotional response to music. Put simply, thinking and feeling in music, as they are in other domains (e.g., Damasio, 2000), are closely linked (Cross, 2005; Meyer, 1956).

Balkwill, Thompson, and Matsunaga (2004) investigated the recognition of emotion by Japanese listeners expressed through Japanese, Western, and Hindustani music. Listeners were sensitive to the intended emotion in the music from all three cultures and judgments of emotion were related to judgments of acoustic cues. For example, ratings of joy were associated with music that was fast in tempo and melodically simple. Acoustic cues, such as tempo, complexity, and loudness, appear to be related systematically to expressed emotion.

Recently, Fritz et al. (2009) presented computer-generated Western piano music excerpts said to convey happy, sad, scared/fearful emotion to a sample of African Mafa people and Western listeners. Stimuli differed with respect to mode, tempo, pitch range, tone density,

and rhythmic regularity. The task was to match a face from the Ekman archive that best fitted the musical stimulus. Both groups recognized the expression of the three basic emotions above chance level; Western listeners recorded significantly higher hit rates than Mafa listeners. The study demonstrates that participants can match auditory and visual Western examples of expressed emotion (but see Russell, 1994) and it typifies culturally unbalanced designs where the musical material used is typical of only one of the groups of participants.

In a more balanced participant-stimulus design, Laukka, Eerola, Thingujam, Yamasaki, and Beller (2010) asked three professional bowed string musicians from each of four different musical cultures to express 11 different emotions through music and then these materials were judged by three groups of listeners from Sweden, India, and Japan. Anger, fear, happiness, and sadness received the highest recognition rates and were recognized across cultures. In keeping with research into acquired expectations and perceptual habits, listeners were more accurate when judging music from familiar cultures. Acoustic analyses showed that cues underpinning basic emotion recognition included event density, pulse, sound level, and energy.

5. Music, motion, and intermodality

Musical phenomena are not purely auditory but intimately related to movement and therefore multimodal (Bharucha et al., 2006; Dissanayake, 2000; Iyer, 2002; Merker, 2000; Toiviainen, Luck, & Thompson, 2010). An analysis of the didjeridu-accompanied dance-song genre *Lirrga* from Wadeye in the Northern Territory of Australia, for example, shows that songs fall into named distinct tempo ranges, which correlate with different meters in the vocal part resulting in rhythmic mode (Barwick, 2002). The significance of rhythmic modes comes from their association with *dance* (Marett, 2005).

Acquired early preferences for beat patterns, and the multimodal nature of musical rhythm learning, have been documented by Phillips-Silver and Trainor (2005). Presented with an ambiguous rhythmic stimulus that could be interpreted in 2 as a march or in 3 as a waltz, half of the 7-month-old infants were bounced on every second beat, and half on every third beat. After the exposure phase, the infants listened longer to (showed preference for) the auditory stimulus with accented beats that matched the beats on which they were bounced compared with the unfamiliar stimulus. Thus, movement to the beat influenced subsequent perception. In a later experiment without movement, infants showed no preference for the duple or triple meter, suggesting that the movement is crucial for this multisensory effect. Experience of body movement therefore plays a role in musical rhythm perception.

Embodiment, tacit in the music action–perception cycle (e.g., Janata & Grafton, 2003), inspired Leman, Desmet, Styns, van Noorden, and Moelants (2009) to record movement velocity patterns of the arm of a musician and of listeners during a performance of Chinese guqin music. Listeners' movement velocity tended to correlate with each other and with movement velocity of the performer's shoulders. The results are interpreted as evidence of shared musical expression and associated ''corporeal intentionality'' (p. 263). Davidson

(2010) compared the learning of music: (a) in a rural Balinese community through both dance and singing; (b) among the Venda people in rural South Africa where dancing and singing are integrated; and (c) in Australian high school students where music, singing, and dancing are learned as separate elements. The results showed that the context in which musical arts activities take place shapes relations of music, singing, and dancing to one another. In Venda, separation between skills of singing and dancing are few; in the Australian sample there is a lack of opportunity to dance rhythms, to feel melodies, and an emphasis in young instrumentalists on achieving fine motor control. The learning experience in the Australian context appears, to Davidson, to be more stressful and less socially communicative—a reminder of the crucial role of context and the body in musical practice and communication.

6. Conclusion

Diverse musical systems and settings are endangered (Huron, 2008). The rapidly spreading monoculture of the West is eradicating the diversity of musical minds. Huron observes that ''even in the western Amazon people listen to Funk Carioca and Christina Aguilera'' (p. 457). Documentation and education are vital (Corn, 2007) and the resulting documentation has potential to broaden and inform current theories of music perception and cognition. As an example, pitch perception is regarded widely as based on frequency ratios (Moore, 2003) and various musical scales are based on logarithmic frequency systems; that is, the invariant interval features are frequency ratios rather than frequency differences characteristic of linear systems (Will & Ellis, 1996). However, Will and Ellis (1994) describe linear frequency structures in Australian Aboriginal music. This observation challenges the assumption that transpositions are based on frequency ratios but the alternative view has had little impact on theories of pitch perception to date. Fortunately, different organizing principles suggested in musicological or acoustic analyses are increasingly accessible and can be investigated experimentally in the context of perception and cognition (Honing, 2006).

The present review has referred to examples of universals and diversity and has emphasized multimodal integration, cross-fertilization between music and language, and the importance of movement in musical perception and action. Discovering new ways of representing pitch, time, timbre, and song will help ensure that theories of music perception and cognition are inclusive and explanatory. Raising awareness of diversity might start the revolution for present generations to recognize, and future generations to learn, diverse musical and cultural forms.

Acknowledgments

Research supported by an Australian Research Council Discovery Project grant (DP0771890). Thanks to the Music Cognition and Action group in MARCS Institute for earlier discussions about these issues.

References

Ayari, M., & McAdams, S. (2003). Aural analysis of Arabic improvised instrumental music (Taqsīm). *Music Perception*, *21*(2), 159–216.

Balkwill, L.-L., Thompson, W. F., & Matsunaga, R. (2004). Recognition of emotion in Japanese, western, and Hindustani music by Japanese listeners. *Japanese Psychological Research*, *46*(4), 337–349.

Barnes, R., & Jones, M. R. (2000). Expectancy attention and time. *Cognitive Psychology*, *41*(3), 254–311.

Barwick, L. (1989). Creative (ir)regularities: The intermeshing of text and melody in performance of Central Australian song. *Australian Aboriginal Studies*, *1989/1*, 12–28.

Barwick, L. (2002). Tempo bands, metre and rhythmic mode in Marri Ngarr ''Church Lirrga'' songs. *Australasian Music Research*, *7*, 67–83.

Bent, T., Bradlow, A. R., & Wright, B. A. (2006). The influence of linguistic experience on the cognitive processing of pitch in speech and nonspeech sounds. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(1), 97–103.

Bharucha, J. J., Curtis, M., & Paroo, K. (2006). Varieties of musical experience. *Cognition*, *100*(1), 131–172.

Bispham, J. (2006). Rhythm in music: What is it? Who has it? And why? *Music Perception*, *24*(2), 125–134.

Blacking, J. (1973). *How musical is man?* London: Faber & Faber.

Blacking, J. (1995). *Music, culture, and experience: Selected papers of John Blacking*. Chicago: The University of Chicago Press.

Bregman, A. (1990). *Auditory scene analysis*. Cambridge, MA: MIT Press.

Brown, S., & Jordania, J. (in press). Universals in the world's musics. *Psychology of Music*, http://pom.sagepub.com/content/early/2011/11/23/0305735611425896 doi:10.1177/0305735611425896.

Burnham, D., & Mattock, K. (2010). Auditory development. In J. G. Bremmer & T. D. Wachs (Eds.) *The Wiley-Blackwell handbook of infant development, Vol. 1: Basic research* (2nd ed., pp. 81–119). Chichester, East Sussex, UK: Wiley-Blackwell.

Carterette, E. C., & Kendall, R. A. (1999). Comparative music perception and cognition. In D. Deutsch (Ed.), *The psychology of music* (pp. 725–791). San Diego: Academic Press.

Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal hierarchies in the music of North India. *Journal of Experimental Psychology: General*, *113*(3), 394–412.

Chandrasekaran, B., Krishnan, A., & Gandour, J. (2009). Relative influence of musical and linguistic experience on early cortical processing of pitch contours. *Brain & Language*, *108*(1), 1–9.

Clayton, M. (1997). Metre and tal in North Indian music. Translated from Le metre et le tal dans la musique de l'Inde du Nord. *Cahiers de Musiques Traditionelles*, *10*, 169–189. Trans. G. Goormaghtigh.

Clayton, M., Herbert, T., & Middleton, R. (Eds.) (2003). *The cultural study of music: A critical introduction*. New York: Routledge.

Clayton, M., Sager, R., & Will, U. (2005). In time with the music: The concept of entrainment and its significance for ethnomusicology. *European Meetings in Ethnomusicology (ESEM Counterpoint)*, *1*, 3–75.

Corn, A. (2007). To sing their voices eternal. *Music Council of Australia Music Forum*, February–April, 22–25.

Cross, I. (2005). Musicality and the human capacity for culture. *Proceedings of the International Symposium on Evolutionary Musicology*. Seewiesen, Germany: Max Planck Institute for Ornithology.

Cuddy, L. L., & Lunney, C. A. (1995). Expectancies generated by melodic intervals: Perceptual judgments of melodic continuity. *Perception & Psychophysics*, *57*(4), 451–462.

Curtis, M. E., & Bharucha, J. J. (2009). Memory and musical expectation for tones in cultural context. *Music Perception*, *26*(4), 365–375.

Curtis, M. E., & Bharucha, J. J. (2010). The minor third communicates sadness in speech, mirroring its use in music. *Emotion*, *10*(3), 335–348.

Damasio, A. (2000). *The feeling of what happens: Body and emotion in the making of consciousness*. London: Vintage.

Davidson, J. (2010). Embodied musical communication: Evidence from singing, dancing and musical instrument playing. In S. M. Demorest, S. J. Morrison, & P. S. Campbell (Eds.), *Proceedings of the 11th international conference on music perception and cognition* (p. 124). Seattle: University of Washington (Abstract).

Deutsch, D., Henthorn, T., & Dolson, M. (2004). Absolute pitch, speech, and tone language: Some experiments and a proposed framework. *Music Perception*, *21*(3), 339–336.

Dissanayake, E. (2000). Antecedents of the temporal arts in early mother-infant interaction. In N. L. Wallin, B. Merker, & S. Brown (Eds.) *The origins of music* (pp. 389–410). Cambridge, MA: MIT Press.

Drake, C., & Bertrand, D. (2001). The quest for universals in temporal processing in music. In R. J. Zatorre & I. Peretz (Eds.), *The biological foundations of music: Annals of the New York Academy of Sciences* (Vol. 930, pp. 17–27). New York: The New York Academy of Sciences.

Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition*, *77*(3), 251–288.

Eerola, T., Himberg, T., Toiviainen, P., & Louhivuori, J. (2006). Perceived complexity of western and African folk melodies by western and African listeners. *Psychology of Music*, *34*(3), 337–371.

Eerola, T., Louhivuori, J., & Lebaka, E. (2009). Expectancy in North Sami yoiks revisited: The role of data-driven and schema-driven knowledge in the formation of melodic expectations. *Musicae Scientiae*, *13*(2), 39–70.

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, *32*(5), 429–492.

Finlay, R. (2011). Salsa. In B. Kerfield (Ed.), *The New Grove Dictionary of Jazz* (2nd ed.). Grove Music Online available at: http://www.oxfordmusiconline.com/subscriber/article/grove/music/J683900. Accessed May 12, 2011.

Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., Friederici, A. D., & Koelsch, S. (2009). Universal recognition of three basic emotions in music. *Current Biology*, *19*(7), 573–576.

Hanna, J. L. (1979). *To dance is human: A theory of nonverbal communication*. Austin: University of Texas Press.

Hannon, E. E., & Trehub, S. E. (2005). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(35), 12639–12643.

Harwood, D. L. (1976). Universals in music: A perspective from cognitive psychology. *Ethnomusicology*, *20*(3), 521–533.

Himberg, T., & Thompson, M. R. (2011). Learning and synchronizing dance movements in South African songs —Cross-cultural motion capture study. *Dance Research*, *29*(2), 305–328.

Honing, H. (2006). On the growing role of observation, formalization and experimental method in musicology. *Empirical Musicology Review*, *1*(1), 1–6.

Huron, D. (2006). *Sweet anticipation: Music and the psychology of expectation*. Cambridge, MA: The MIT Press.

Huron, D. (2008). Lost in music. *Nature*, *453*, 456–457.

Iversen, J. R., Patel, A. D., & Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. *Journal of the Acoustical Society of America*, *124*(4), 2263–2271.

Iyer, V. S. (1998). *Microstructures of feel, macrostructures of sound: Embodied cognition in West African and African-American musics*. Unpublished doctoral dissertation, Technology and the Arts, University of California, Berkeley.

Iyer, V. (2002). Embodied mind, situated cognition, and expressive microtiming in African-American music. *Music Perception*, *19*(3), 387–414.

Janata, P., & Grafton, S. T. (2003). Swinging in the brain: Shared neural substrates for behaviors related to sequencing and music. *Nature Neuroscience*, *6*(7), 682–687.

Jones, M. R. (2004). Attention and timing. In J. G. Neuhoff (Ed.), *Ecological psychoacoustics* (pp. 49–87). San Diego, CA: Elsevier Academic Press.

Juslin, P. N., & Västfjäll, D. (2008). Emotional response to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*(5), 559–621.

Kessler, E. J., Hansen, C., & Shepard, R. N. (1984). Tonal schemata in the perception of music in Bali and the West. *Music Perception*, *2*(2), 131–165.

Kirschner, S., Ilari, B., & Tomasello, M. (2010). Joint drumming in Brazilian and German preschool children: Synchronization and prosocial effects. In S. M. Demorest, S. J. Morrison, & P. S. Campbell (Eds.), *Proceedings of the 11th international conference on music perception and cognition* (pp. 753–754). Seattle: University of Washington (Abstract).

Krumhansl, C. L., Toivanen, P., Eerola, T., Toiviainen, P., Järvinen, T., & Louhivuori, J. (2000). Cross-cultural music cognition: Cognitive methodology applied to North Sami yoiks. *Cognition*, *76*(1), 13–58.

Large, E., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*(1), 119–159.

Laukka, P., Eerola, T., Thingujam, N. S., Yamasaki, T., & Beller, G. (2010). Universal and culture-specific factors in the recognition and performance of musical emotions. In S. M. Demorest, S. J. Morrison, & P. S. Campbell (Eds.), *Proceedings of the 11th international conference on music perception and cognition* (p. 97). Seattle: University of Washington (Abstract).

Leman, M., Desmet, F., Styns, F., van Noorden, L., & Moelants, D. (2009). Sharing musical expression through embodied listening: A case study based on Chinese guqin music. *Music Perception*, *26*(3), 263–278.

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

Magill, J. M., & Pressing, J. L. (1997). Asymmetric cognitive clock structures in West African rhythms. *Music Perception*, *15*(2), 189–222.

Marett, A. (2005). *Songs, dreamings, and ghosts: The Wangga of North Australia*. Middletown, CT: Wesleyan University Press.

McAdams, S. (1989). Psychological constraints on form-bearing dimensions in music. *Contemporary Music Review*, *4*(1), 181–198.

McMullen, E., & Saffran, J. R. (2004). Music and language: A developmental comparison. *Music Perception*, *21*(3), 289–311.

Merker, B. (2000). Synchronous chorusing and human origins. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 315–327). Cambridge, MA: MIT Press.

Merker, B. (2005). The conformal motive in birdsong, music, and language: An introduction. *Annals of the New York Academy of Sciences*, *1060*, 17–28.

Merker, B. H., Madison, G. S., & Eckerdal, P. (2009). On the role and origin of isochrony in human rhythmic entrainment. *Cortex*, *45*(1), 4–17.

Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.

Meyer, L. B. (1960). Universalism and relativism in the study of ethnic music. *Ethnomusicology*, *4*(2), 49–54.

Moore, B. C. J. (2003). *An introduction to the psychology of hearing* (5th ed.). Boston: Academic Press.

Nam, U. (1998). Pitch distributions in Korean court music: Evidence consistent with tonal hierarchies. *Music Perception*, *16*(2), 243–247.

Nan, Y., Knösche, T. R., & Friederici, A. D. (2006). The perception of musical phrase structure: A cross-cultural ERP study. *Brain Research*, *1094*(1), 179–191.

Nettl, B. (2005). *The study of ethnomusicology: Thirty-one issues and concepts*. Urbana: University of Illinois Press.

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *The Journal of Neuroscience*, *31*(28), 10234–10240.

Patel, A. D. (2006). Musical rhythm, linguistic rhythm, and human evolution. *Music Perception*, *24*(1), 99–104.

Patel, A. D., & Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition*, *87*(1), B35–B45.

Patel, A. D., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Current Biology*, *19*(10), 1–4.

Petocz, A., Keller, P., & Stevens, C. (2008). Auditory warnings, signal-referent relations and natural indicators: Re-thinking theory and application. *Journal of Experimental Psychology: Applied*, *14*(2), 165–178.

Pfordresher, P. Q., & Brown, S. (2009). Enhanced production and perception of musical pitch in tone language speakers. *Attention, Perception, & Psychophysics*, *71*(6), 1385–1398.

Phillips-Silver, J., Aktipis, C. A., & Bryant, G. A. (2010). The ecology of entrainment: Foundations of coordinated rhythmic movement. *Music Perception*, *28*(1), 3–14.

Phillips-Silver, J., & Keller, P. E. (2012). Searching for roots of entrainment and joint action in early musical interactions. *Frontiers in Human Neuroscience*, *6*, 1–11.

Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science*, *3078*, 1430.

Povel, D.-J. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(1), 3–18.

Pressing, J. (2002). Black Atlantic rhythm: Its computational and transcultural foundations. *Music Perception*, *19*(3), 285–310.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin*, *115*(1), 102–141.

Schellenberg, E. G., & Trehub, S. E. (2008). Is there an Asian advantage for pitch memory? *Music Perception*, *25*(3), 241–252.

Schmuckler, M. A. (1989). Expectation in music: Investigation of melodic and harmonic processes. *Music Perception*, *7*(2), 109–150.

Soley, G., & Hannon, E. E. (2010). Infants prefer the musical meter of their own culture: A cross-cultural comparison. *Developmental Psychology*, *46*(1), 286–292.

Stevens, C. J., Keller, P. E., & Tyler, M. D. (in press) Tonal language background and detecting pitch contour in spoken and musical items. *Psychology of Music*, http://pom.sagepub.com/content/early/2011/08/13/0305735611415749 doi:10.1177/0305735611415749.

Stobart, H., & Cross, I. (2000). The Andean anacrusis? Rhythmic structure and perception in Easter songs of Northern Potosí, Bolivia. *British Journal of Ethnomusicology*, *9*(ii), 63–94.

Tenzer, M. (2006). Oleg Tumulilingan. Layers of time and melody in Balinese music. In M. Tenzer (Ed.), *Analytical studies in World music* (pp. 205–236). Oxford, England: Oxford University Press.

Thompson, W. F., & Schellenberg, E. G. (2006). Listening to music. In R. Colwell (Ed.), *MENC handbook of musical cognition and development* (pp. 72–123). Oxford, England: Oxford University Press.

Titon, J. T. (Ed.) (1992). *Worlds of music: An introduction to the music of the world's peoples*. New York: Schirmer Books.

Toiviainen, P., & Eerola, T. (2003). Where is the beat? Comparison of Finnish and South African listeners. In R. Kopiez, A. C. Lehmann, I. Wolther, & C. Wolf (Eds.), *Proceedings of the 5th triennial ESCOM conference* (pp. 501–504) Germany: Hanover University of Music and Drama.

Toiviainen, P., Luck, G., & Thompson, M. R. (2010). Embodied meter: Hierarchical eigenmodes in music-induced movement. *Music Perception*, *28*(1), 59–70.

Turpin, M. (2007a). Artfully hidden: Text and rhythm in a central Australian Aboriginal song series. *Musicology Australia*, *XXIX*, 93–108.

Turpin, M. (2007b). The poetics of central Australian song. *Australian Aboriginal Studies*, *2007/2*, 100–115.

Velasco, M. J., & Large, E. W. (in press). Pulse detection in syncopating rhythms using neural oscillators. In A. Klapuri & C. Leider (Eds.), *Proceedings of the 12th annual conference of the International Society for Music Information Retrieval* (pp. 185–190). Miami, FL: University of Miami.

Vos, P. G., & Troost, J. M. (1989). Ascending and descending melodic intervals: Statistical findings and their perceptual relevance. *Music Perception*, *6*(4), 383–396.

Will, U., & Ellis, C. (1994). Evidence for linear transposition in Australian Western Desert vocal music. *Musicology Australia*, *XVII*(1), 2–12.

Will, U., & Ellis, C. (1996). A re-analyzed Australian Western Desert song: Frequency performance and interval structure. *Ethnomusicology*, *40*(2), 187–222.

Wong, P. C. M., Roy, A. K., & Margulis, E. H. (2009). Bimusicalism: The implicit dual enculturation of cognitive and affective systems. *Music Perception*, *27*(2), 81–88.

Yip, M. J. W. (2002). *Tone*. Cambridge, England: Cambridge University Press.

Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(13), 5768–5773.