

ESSAY

The neural roots of music

Laurel Trainor explains how the emotional power of music depends on the structure of the ear, and on our basic encoding of information.

In search of an original voice, the dominant composers of the mid-twentieth century — Arnold Schoenberg, Pierre Boulez and their disciples — rejected the tonal and rhythmic forms of the past. They adhered to rigorous compositional techniques such as the serial tone-row method — in which all notes of the chromatic scale occur equally often in a repeating row — banishing tonality. Some powerful compositions were written in the serial style, but few are played regularly today. Asked in 1999 why this might be, Boulez responded: “Well, perhaps we did not take sufficiently into account the way music is perceived by the listener.”

Understanding the structure and development of our auditory pathways, and how experience modifies them, may inform us about why some musical experiments are successful, and others are not. Why, for instance, the melodies of Gustav Mahler, the driving rhythms of Igor Stravinsky or the dissonances of John Adams make these modern composers popular today, whereas the music of some others, such as Luigi Dallapiccola and Luigi Nono, is rarely heard.

Music is conceived by our brains, played through our bodies, perceived through our sensory organs and then interpreted by our brains. Thus it is subject both to general constraints of our neural system and to specific constraints of our auditory processing capacities.

During childhood, each of the billions of neurons in the human auditory system forms thousands of connections to other neurons, creating neural networks. Genes control the characteristics of neural circuits, developmental waves of neuronal and synaptic proliferation, and the later pruning of neural connections to form efficient circuits for processing sound. Experience also profoundly affects the neural connections formed. Studies show that rats raised in environments containing only white noise with no pitch or rhythm are unable to recognize everyday sounds, and are greatly impaired even in their ability to discriminate different pitches¹.

Certain sounds elicit specific, powerful emotions in people, presumably a testament to the evolutionary heritage of our auditory systems. Low, loud, dissonant sounds evoke fear; rapid, higher, consonant sounds evoke friendliness or joy. Mothers around the world talk and sing to infants using a cooing tone of voice and higher pitch than when interacting with adults. Infants prefer these higher-pitched vocalizations and

mothers sing in different styles to help prelinguistic infants regulate their emotional state. Across cultures, songs sung while playing with babies are fast, high and contain exaggerated rhythmic accents; lullabies are lower, slower and softer.

Talking to people of all ages, we use falling pitches to express comfort; relatively flat, high pitches to express fear; and large bell-shaped pitch contours to express joy and surprise. Hearing music with an unfamiliar structure, listeners base their emotional reactions largely on such sound features.

Music is built on general, universal features of human sound processing that have deep evolutionary roots. It also incorporates rhythmic, melodic and harmonic structure. Musical structures and styles vary enormously across cultures, and change as continually as languages, yet our biology constrains the possibilities.

Rhythm is a dancer

Musical rhythm may have its origins in the motor rhythms controlling locomotion, breathing and heart rate. Babies receive correlated sound and movement input as parents rock them while singing. This and other early experience encourages movement and auditory representations to wire together in the brain.

Although music makes us want to move to the beat, movement evolved first and there are multi-sensory connections in the brain between motor and auditory areas. Therefore it should also be the case that how we move affects how we interpret rhythms. We have shown this with a repeating 6-beat rhythm pattern with no accents that can be perceived as two groups of three beats (as in a waltz) or as three groups of two beats (as in a march). Adults and infants who bounce up and down on every second beat report hearing — or in the case of the infants prefer — a march. Those bouncing on every third beat hear a waltz².

The evolutionarily ancient vestibular system for balance plays a crucial role in the interaction between movement and the perception of musical rhythm, indicating that music and dance could have evolved together. Stimulating someone's vestibular nerve alternately on the left and right sides gives a sensation that the head is moving from side to side. Such stimulation alone, on either every second beat or on every third beat of an unaccented 6-beat rhythm pattern, biases judgement of whether

the music is heard as a march or a waltz³.

Western music, from pop to classical, tends to use simple rhythmic structures. Folk music from many other traditions uses complex structures, for example pitting groups of 7 and 11 beats against each other. Young infants can perceive complex rhythmic structure, but they lose this ability before they are a year old if not exposed to such rhythms.

Our capacity for processing rhythmic complexity is likely to underlie the comparative success of rhythmic experimentation over pitch experimentation in the twentieth century. It might explain why rhythmic structures from many traditions around the world have been successfully incorporated into music, why audiences embrace the jazz rhythms of composers such as George Gershwin, and why people find the rhythmic richness of composers such as Stravinsky, Béla Bartók, Jennifer Higdon and film composer Danny Elfman so compelling.

Pitch invasion

The use of pitch in music reflects the constraints of auditory mechanisms for identifying objects and separating sound sources. These constraints are evident in the near-universal use of consonance and dissonance as an organizing principle; in the use of scales comprising a small set of pitch categories that repeat at octave intervals, reducing the amount the listener has to remember; and in the use of at least two different-sized intervals in scales, allowing the emergence of different relationships between tone pairs and tonal functions such as ‘tonic’ and ‘dominant’ in Western diatonic scales.

Sensory dissonance arises from how the basilar membrane vibrates in the cochlea of the inner ear, and from the firing patterns of auditory nerve fibres that this movement activates. The basilar membrane acts as a sort of Fourier analyser. It reacts to each frequency component of a sound, with the point of maximal





vibration near one end for low frequencies and near the other for high frequencies. Two simultaneous frequencies that are less than a critical bandwidth apart cause vibration patterns that interact on the membrane.

This is why it is difficult to hear individual tones in chord clusters with small pitch distances between adjacent notes.

The perceived pitch of a sound corresponds to its energy at integer multiples — harmonics — of a fundamental frequency. Two sounds containing harmonics within critical bandwidths make interference patterns on the basilar membrane, and produce a sense of dissonance. The frequency content of tones is also processed according to when neurons fire, and consonant and dissonant stimuli cause different types of firing patterns in auditory nerves.

Emotions arise in part through the ebb and flow of tension in music. Alternation between consonance and dissonance is a powerful device in this regard. Dissonance can be very beautiful, and resolution to consonance especially poignant. Composers can choose to ignore the fundamental physiological power of the consonance–dissonance relation and the information-processing power of discrete pitch and unequal interval sizes in scales. But by doing so, they create music that demands more of the listener because it lacks some of the most powerful physiological organizing principles of our nervous system.

Perception also depends on experience. During development, infants and children learn the pitch organization of their culture's music types and thereafter process music through the filter of this knowledge. Even musically untrained Western listeners acquire implicit knowledge of the Western major scale. They readily detect a wrong note that goes outside the scale on which a melody is based. They have

considerably more trouble detecting changes within the scale because these do not violate their implicit knowledge of which notes belong in the key. Infants under one, on the other hand, do not yet process music according to particular scales, and notice changes that violate major scale structure and changes that do not.

Experience counts

Harmonic structure (sequences of chords that follow each other according to syntactic rules) dominates Western music, but is relatively rare across musical systems. Without specific musical training, sensitivity to harmonic structure emerges in children only after about 5 years of age⁴.

Scale and harmonic structures depend on learning. By contrast with sensory consonance and dissonance, there is more flexibility in how they are perceived. They also show greater diversity across the world's musics. For example, different intervals are used in the Western major and minor scales, the pentatonic scale, and the many melodic modes (*rāgas*) used in Indian classical music. Whereas many traditions, such as *rāga* improvisations, employ a drone, or single pitch, over which the melody is played, fully developed harmonic syntax as in Western music is very rare. Our ability to learn different scales and harmonic structures gives composers considerable flexibility for experimentation that audiences can perceive and appreciate.

To recap: the spectral and temporal organization of music — its rhythm and pitch — derives from our biology. Neural constraints dictate that some musical structures are easier

to perceive and learn, giving rise to some near-universal features of music. Music is difficult to process when consonance and dissonance do not anchor the ebb and flow of tension, and when all pitches are equally prominent. Such music has no point from which to interpret pitch intervals. For many listeners, this level of difficulty is not enjoyable. Equally, music that is too simple and predictable can be boring.

The flexibility of our auditory system and its dependence on learning enables us to invent different musical structures, and allows musical tastes to change with familiarity and experience. There are notable examples of audience revolts at premier performances of works that seem tame to future generations — Stravinski's ballet *The Rite of Spring* caused a riot and Beethoven's third symphony was incomprehensible to reviewers. What has not changed recently is our evolutionary inheritance, the structure of our sensory organs, our basic encoding of information and our visceral responses to features of sound that unleash the emotional power of music in our lives. ■

Laurel Trainor is director of the McMaster Institute for Music and the Mind, and the Auditory Development Lab at McMaster University, 1280 Main Street West, Hamilton, Ontario L854L8, Canada.

1. Chang, E. F. & Merzenich, M. M. *Science* **300**, 498–502 (2003).
2. Phillips-Silver, J. & Trainor, L. J. *Science* **308**, 1430 (2005).
3. Trainor, L. J., Gao, X., Lei, J., Lehtovarara, K. & Harris, L. R. *Cortex* (in the press).
4. Trainor, L. J. *Dev. Psychobiol.* **46**, 262–278 (2005).

See <http://tinyurl.com/6e9mt8> for further reading.