Perspectives on the Emotional Response to Consonance and Dissonance in Music
Thalia Vrantsidis

Abstract
In music, the tendency for certain combinations of notes to sound either pleasant (i.e., consonant) or unpleasant (i.e., dissonant) seems to be innate, universal, and uniquely human—provoking the question of why these intervals are able to trigger these emotional reactions. Evidence will be reviewed here to explore several aspects of this question: (1) which neural circuits are involved; (2) what acoustic properties lead to emotional reactions; and (3) why these preferences might have evolved. Brain imaging research shows that many emotion-processing areas respond differentially to consonant and dissonant music. In addition, lesion studies have found specialized neural circuitry involved in emotional responses to music. Studies of people with amusia, a pitch processing disorder, have identified the acoustic properties that determine a sound’s pleasantness—specifically, the harmonic series, rather than beating as previously assumed, distinguishes consonant from dissonant intervals. Consonant intervals are also the ones most prevalent in the harmonics of the human voice. This may explain why humans show a preference for consonant intervals.

“Ah, music,” he said, wiping his eyes. “A magic beyond all we do here!”
– Dumbledore, in J.K. Rowling’s Harry Potter and the Sorcerer’s Stone (1997, p. 95)

Music has the power to make people feel. It inspires, it soothes, it makes us happy, and it makes us cry. Music is used to emotionally enrich the vast experiences of life, from milestone celebrations and sporting events, right down to personal listening (McDermott & Hauser, 2005). How is it that a pattern of notes can inspire such reactions? Part of the mystery of music is that there does not seem to be an obvious evolutionary benefit to it. Furthermore, it is not clear why certain patterns of notes produce certain emotional responses: why does Mozart sound beautiful, but a child smashing the keys of a piano sound terrible? Two of the most basic musical aspects that evoke emotional responses are consonance and dissonance, which define combinations of notes that sound either pleasant or unpleasant, respectively. There is widespread agreement that certain intervals sound highly pleasant, for example, a perfect 5th (defined by a distance of 7 semitones), while others are regarded as highly unpleasant, especially the tritone (defined by a distance of 6 semitones) (Krumhansl, 1990; Malmberg, 1918; McDermott, Lehr, & Oxenham, 2010). This paper will examine the neuroscientific research to explore the neural underpinnings of the emotional reaction to consonance and dissonance, and discuss the acoustical properties of the sounds which trigger these responses. In addition, this research will be connected to a theory of why emotional responses to consonance and dissonance might have evolved.

Preferences for consonant over dissonant intervals are a promising starting point for understanding human emotional reactions to music, as they appear to have an evolutionary basis that is unique to humans. Masataka (2006) showed that human infants as young
as two days old show preferences for consonant over dissonant music, indicating the preference is intrinsic in nature. This occurred equally in children of deaf and hearing parents, suggesting that these preferences are formed independently of pre-natal exposure to music, which presumably would differ with deaf and hearing parents. In addition, cross-cultural studies have shown similar preferences in many non-Western cultures (Butler & Daston, 1968; Fritz et al., 2009) although exceptions have been found in Indian listeners, perhaps due to the increased exposure to dissonance as used in Indian classical music (Maher, 1976). These results suggest that the human preference for consonance may be innate, although alterable by experience, and thus have an evolutionary basis. Furthermore, although non-human species such as birds and monkeys can be trained to differentiate between consonant and dissonant sounds (Hulse, Bernard, & Braaten, 1995; Izumi, 2000), humans are the only species known to show a preference between interval types (McDermott & Hauser, 2004). The emotional reactions to consonance and dissonance are an innate and uniquely human response, and can serve as a window into the human experience of music.

Review of Neuroscientific Literature

While the neuroscience of musical emotions is still in its infancy, it has made considerable strides in understanding the neural underpinnings of the emotional response to consonance and dissonance. Notably, it has been found that the brain regions which respond preferentially to either consonant or dissonant music are also involved in processing other pleasant or unpleasant emotional stimuli. For example, a study conducted by Blood, Zatorre, Bermudez, and Evans (1999) utilized positron emission tomography (PET) to investigate brain responses to music with varying degrees of dissonance. Dissonance was positively correlated with activation in the parahippocampal gyrus, an area previously shown to react to pictures with a negative emotional valence (Lane et al., 1997). Conversely, consonance was positively correlated with activation of the subgenual cingulate, frontopolar and orbitofrontal cortices. These regions have also been shown to be involved in emotional processing; for example, lesions of the subgenual cingulate and other ventral medial prefrontal cortex regions in monkeys have been shown to impair recognition of emotional expressions (Hornak, Rolls, & Wade, 1996).

A study utilizing functional magnetic resonance imaging (fMRI) (Koelsch, Fritz, Müller, & Friederici, 2006) compared responses to joyful (i.e., mainly consonant) instrumental dance tunes and dissonant versions of those same tunes. The consonant music increased activation in the ventral striatum, containing the nucleus accumbens, which has been proposed as one of the main structures involved in the motivational and rewarding effects of positive stimuli (Cardinal, Parkinson, Hall, & Everitt, 2002; Salamone, Correa, Mingote, & Weber, 2005). The ventral striatum has also been shown to activate during the highly pleasurable experience of “chills” during music-listening (Blood & Zatorre, 2001). The dissonant music increased activation in the amygdala, hippocampus, parahippocampal gyrus, and temporal poles—areas that are highly interconnected and form a core network involved in reactions to aversive stimuli (see Koelsch et al., 2006, for a review). It is important to note that the involvement of the amygdala during emotional processing of music is still debated. It does not seem to be confined to aversive stimuli: Ball et al. (2007) showed no significant differences in amygdala activation between dissonant and consonant music, whereas Blood and Zatorre (2001) showed that the amygdala was involved during pleasurable music-induced chills.

Heart rate monitors and electroencephalography (EEG) have also been used to measure responses to consonant and dissonant music. Sammler, Grigutsch, Fritz, and Koelsch, (2007) found that, similar to unpleasant pictures, sounds, and films, dissonant music (relative to consonant music) led to a decreased heart rate, implying common underlying neural mechanisms.
in the processing of these stimulus types. They also found an increase in frontal-midline (Fm) theta power for consonant music, indicating greater coherence of neural firing in certain population of neurons. This is interpreted as reflecting activation in the anterior cingulate cortex (ACC), which has been implicated elsewhere in processing pleasant music (Blood & Zatorre 2001). The ACC’s role in emotion is also supported by its inputs from other key emotion-related areas, such as the amygdala, hippocampus, and parahippocampal gyrus (Devinsky, Morrell, & Vogt, 1995).

Despite the differences in brain activation patterns across these studies, it is clear that the major brain areas responsible for emotional processing are also being recruited for responding to consonant and dissonant music. In addition to these areas, Griffiths, Warren, Dean, and Howard (2004) may have discovered a brain area specific to musical emotions, rather than emotions in general. They reported a case study of a patient who, several months after a stroke, had selectively lost all emotional responses to music, while retaining his capacities for musical perception and emotional responses to non-musical stimuli. fMRI scans revealed damages mainly to the left insula, which also extended into the left frontal lobe and the left amygdala. This damage overlaps with areas shown to be active in normal emotional processing of music (Blood & Zatorre, 2001), suggesting that some of these damaged areas might form a specialized neural substrate for musical emotions, separate from music perception and non-musical emotional responses.

**Acoustical Basis of Consonance and Dissonance**

Over the centuries, various theories have been advanced for what properties of consonant versus dissonant sounds might trigger the preference for consonant intervals (see Cousineau, McDermott, & Peretz, 2012, for a review). The most common theory is based on “beating”, a phenomenon in which two waves of different frequencies are combined, and its resulting sound wave is experienced as having an unpleasant sound quality known as roughness. Beating was generally thought to be more prevalent in dissonant than consonant intervals and was thus used to explain the dislike of dissonant intervals (although see Cousineau, McDermott, & Peretz, 2012, for evidence that beating is generally equally prevalent in both types of intervals). A recent study tested this theory using people with amusia, a disorder characterised by difficulty in pitch processing (Cousineau, McDermott, & Peretz, 2012). Amusics showed no difference in liking of consonant and dissonant intervals; however, they still disliked notes with beating. Furthermore, McDermott, Lehr, and Oxenham (2010) found that, among individuals with normal auditory functioning, dislike of beating was not correlated with dislike of dissonant notes. These studies thus demonstrate that beating is not able to explain the preference of consonant over dissonant intervals.

An alternative explanation exists based on the harmonic series, specifically the harmonics series present in the human voice. Naturally produced pitches do not contain a single frequency; instead, they are composed of a fundamental frequency accompanied by several sound waves with frequencies that are integer multiples of the fundamental frequency (i.e., harmonics; Cousineau, McDermott, & Peretz, 2012). Generally, the fundamental frequency is loudest (the wave has the largest amplitude), with each higher harmonic decreasing in volume (Smith, 2003). However, in ordinary human speech, certain harmonics tend to be louder, especially the third and fourth harmonics, due to properties of the human vocal tract (Schwartz, Howe, & Purves, 2003). Schwartz et al. describe how this creates additional amplitude peaks, at certain frequencies particular to the human voice. They then show that the relationship of the peaks to the fundamental frequency corresponds to the frequency ratios of each of the 12 notes of the chromatic scale compared to the tonic (the first note). Furthermore, the relative amplitude of the peaks can be used to predict the consonance of each interval.
in the chromatic scale, so that for larger peaks the corresponding interval is more consonant. Thus, the particular sounds that are considered consonant or dissonant seem to stem from the unique harmonic properties of the human voice.

Based on these results, Schwartz and colleagues (2003) posited that the preference for consonance might have developed due to its biological utility. One of the primary tasks of the auditory system is to use sound information to identify biologically relevant stimuli. For humans, one of the most pervasive and biologically relevant sources of sound is human speech. However, the identity of a sound source is often ambiguous, so the brain relies on probabilistic rules to help identify the source—in this case, the fact that a certain pattern of harmonics is predictive of the human voice. Schwartz et al. thus proposed that a positive emotional reaction to the human voice might have generalized onto anything which resembled that harmonic pattern, leading to a preference for consonant intervals. This theory can explain the near universality of consonance rankings, as the same pattern of harmonics is present in all human speech, regardless of language (Schwartz et al., 2003). It also explains why other species do not show consonance preferences, since this harmonic pattern is unique to the human voice. Note that this theory does not specify whether the preferences are learned or are innate; however, the aforementioned developmental research would support an innate preference, for which Schwartz et al.'s theory could describe an evolutionary benefit.

**Conclusion**

Neuroimaging studies have broadened our understanding of the neural underpinnings involved in emotional responses to consonance and dissonance. These emotional responses generally involve the same neural circuitry that reacts to other pleasant and unpleasant stimuli. For consonance, this includes the subgenual and anterior cingulate cortex, the frontopolar and orbitofrontal cortex, the ventral striatum, and potentially the amygdala. For dissonance, it involves the parahippocampal gyrus, hippocampus, and temporal poles, and amygdala. The case study of a patient with lesions to the insula, amygdala, and frontal lobe implies that, within these regions, there is specific area for processing musical emotions in particular.

Furthermore, studies of amusic individuals have revealed that the preference for consonance over dissonance does not, as previously thought, stem from the acoustical phenomenon of beating. Instead, statistical analyses suggest that it is specifically based on the similarity of consonant intervals to the harmonics of the human voice. This led Schwartz et al. (2003) to propose that a preference for the human voice provides a plausible explanation for the widespread and uniquely human preference for consonance over dissonance.

**Future Research**

One outstanding question which could be addressed by future studies is the degree to which innate emotional reactions to consonance and dissonance can be altered through experience, as was suggested by the finding that Indian listeners did not show the common pattern of consonance preferences (Maher, 1976). Future research could confirm this explanation by experimentally manipulating participants’ exposure to consonance and dissonance.

Mull (1957) addressed this to some extent, showing that repeated exposure to dissonant music did in fact increase participants’ liking of it; however, it is possible that listeners did not actually enjoy the dissonance more, but were simply more able to focus on and appreciate the other aspects of the music. The use of more controlled stimuli along with neuroimaging could address this question more thoroughly. As a possible experiment, repeated listening over several days or weeks could be used to induce familiarity to either dissonant music or consonant music. Then single dissonant or consonant intervals could be played for the
participants, rather than full songs, while they underwent fMRI and rated how much they liked the intervals. If familiarity with dissonant music did indeed increase participants’ liking of isolated dissonant intervals, this would be shown in the ratings; also, the brain scans for those exposed to dissonant music should show less activation in the areas associated with unpleasantness in previous studies. This study could thus demonstrate the extent to which the dislike of dissonance can be altered as a function of experience, and help to reconcile people’s innate preferences for consonance with the cross-cultural variation found by Maher (1976). In addition to showing how the usual pattern of consonance preferences can be altered, these results could also highlight the potential importance of experience in strengthening the commonly-found consonance preferences, through a lifetime of exposure to consonant music and the human voice. Thus, although there does seem to be an innate basis for emotional reactions to consonance and dissonance, there is a role for both nature and nurture in a full explanation of this most basic aspect of the musical experience.

References


Footnotes

1 A semitone is the smallest interval between notes used in ordinary musical scales (equivalent to the interval between a white key and an adjacent black key on a piano, or between a C and C#).