Behavioral/Systems/Cognitive

Neural Correlates of Consonance, Dissonance, and the Hierarchy of Musical Pitch in the Human Brainstem

Gavin M. Bidelman and Ananthanarayan Krishnan

Department of Speech Language, Hearing Sciences, Purdue University, West Lafayette, Indiana 47907-2038

Consonant and dissonant pitch relationships in music provide the foundation of melody and harmony, the building blocks of Western tonal music. We hypothesized that phase-locked neural activity within the brainstem may preserve information relevant to these important perceptual attributes of music. To this end, we measured brainstem frequency-following responses (FFRs) from nonmusicians in response to the dichotic presentation of nine musical intervals that varied in their degree of consonance and dissonance. Neural pitch salience was computed for each response using temporally based autocorrelation and harmonic pitch sieve analyses. Brainstem responses to consonant intervals were more robust and yielded stronger pitch salience than those to dissonant intervals. In addition, the ordering of neural pitch salience across musical intervals followed the hierarchical arrangement of pitch stipulated by Western music theory. Finally, pitch salience derived from neural data showed high correspondence with behavioral consonance judgments ($r = 0.81$). These results suggest that brainstem neural mechanisms mediating pitch processing show preferential encoding of consonant musical relationships and, furthermore, preserve the hierarchical pitch relationships found in music, even for individuals without formal musical training. We infer that the basic pitch relationships governing music may be rooted in low-level sensory processing and that an encoding scheme that favors consonant pitch relationships may be one reason why such intervals are preferred behaviorally.

Introduction

Relationships between musical pitches are described as either consonant, associated with pleasantness and stability, or dissonant, associated with unpleasantness and instability. Given their anchor-like function in musical contexts, consonant pitch relationships occur more frequently in tonal music than dissonant relationships (Budge, 1943; Vos and Troost, 1989; Huron, 1991). Indeed, behavioral studies reveal that listeners prefer these intervals to their dissonant counterparts (Plomp and Levelt, 1965; Kameoka and Kuriyagawa, 1969b) and assign them higher status in hierarchical ranking (Malmberg, 1918; Krumhansl, 1990; Itoh et al., 2003a). In fact, preference for consonance is observed early in life, well before an infant is exposed to culturally specific music (Trainor et al., 2002; Hannon and Trainor, 2007). It is this hierarchical arrangement of pitch that largely contributes to the sense of a musical key and pitch structure in Western tonal music (Rameau, 1722).

Neural correlates of musical scale pitch hierarchy have been identified at the cortical level in humans using both event-related potentials (Krohn et al., 2007) and functional imaging (Minati et al., 2008). Importantly, these properties are encoded even in adults who lack musical training (Besson and Faita, 1995; Koelsch et al., 2000), leading some investigators to suggest that the bias for these pleasant-sounding pitch intervals may be a universal trait of music cognition (Fritz et al., 2009). These studies show that brain activity is especially sensitive to the pitch relationships found in music and, moreover, that it is enhanced when processing consonant, relative to dissonant, intervals. Animal studies corroborate these findings revealing that the magnitude of phase-locked activity in auditory nerve (Tramo et al., 2001), inferior colliculus (McKinney et al., 2001), and primary auditory cortex (Fishman et al., 2001) correlate well with the perceived consonance/dissonance of musical intervals. Together, these findings offer evidence that the preference for consonant musical relationships may be rooted in the fundamental processing and constraints of the auditory system (Tramo et al., 2001; Zatorre and McGill, 2005; Trainor, 2008). To this end, we hypothesized that preattentive, sensory-level processing in humans may, in part, account for the perceived consonance of musical pitch relationships.

As a window into the early stages of subcortical pitch processing, we use the scalp-recorded frequency-following response (FFR). The FFR reflects sustained phase-locked activity from a population of neural elements within the rostral brainstem and is characterized by a periodic waveform that follows the individual cycles of the stimulus. It is thought that the inferior colliculus (IC), a prominent auditory relay within the midbrain, is the putative neural generator of the FFR (Smith et al., 1975; Galbraith et al., 2000). Use of the FFR has revealed effects of experience-dependent plasticity in response to behaviorally relevant stimuli including speech (Krishnan and Gandour, 2009) and music (Musacchia et al., 2007; Wong et al., 2007). Lee et al. (2009) recently compared brainstem responses in musicians and non-

Table 1. Musical interval stimuli used to evoke brainstem responses

Interval	Musical pitches	No. of semitones	Ratio of fundamentals	Frequency components (Hz)
Unison (Un)	**A3, A3**	**0**	**1:1**	Note 1: ***220, 440, 660, 880, 1100, 1320***
				Note 2: ***220, 440, 660, 880, 1100, 1320***
Minor 2nd (m2)	A3, B♭3	1	16:15	220, 440, 660, 880, 1100, 1320
				235, 470, 705, 940, 1175, 1410
Major 3rd (M3)	**A3, C#4**	**4**	**5:4**	**220, 440, 660, 880, *1100*, 1320**
				275, 550, 825, *1100*, 1375, 1650
Perfect 4th (P4)	**A3, D4**	**5**	**4:3**	**220, 440, 660, 880, 1100, 1320**
				293, 586, 879, 1172, 1465, 1758
Tritone (TT)	A3, D#4	6	45:32	220, 440, 660, 880, 1100, 1320
				309, 618, 927, 1236, 1545, 1854
Perfect 5th (P5)	**A3, E4**	**7**	**3:2**	**220, 440, *660*, 880, 1100, *1320***
				330, *660*, 990, *1320*, 1650, 1980
Major 6th (M6)	**A3, F#4**	**9**	**5:3**	**220, 440, 660, 880, 1100, 1320**
				367, 734, 1101, 1468, 1835, 2202
Major 7th (M7)	A3, G#4	11	15:8	220, 440, 660, 880, 1100, 1320
				413, 826, 1239, 1652, 2065, 2478
Octave (Oct)	**A3, A4**	**12**	**2:1**	**220, *440*, 660, *880*, 1100, *1320***
				***440, 880, 1320*, 1760, 2200, 2640**

Italic values represent frequency components shared between both notes in a given dyad. Bold intervals are consonant; lightface intervals, dissonant. Harmonics of individual notes were calculated based on the ratio of their fundamental frequencies (i.e., just intonation).

musicians to two musical intervals, a consonant major sixth and a dissonant minor seventh, and found response enhancements for musicians to harmonics of the upper tone in each interval. However, overall, responses were similar between consonant and dissonant conditions, and given the limited stimulus set (two intervals), no conclusions could be drawn regarding the possible differential encoding of consonance and dissonance across musical intervals. Expanding on these results, we examine temporal properties of nonmusicians' FFR in response to nine musical intervals to determine whether consonance, dissonance, and the hierarchy of musical pitch arise from basic sensory-level processing inherent to the auditory system.

Materials and Methods

Participants. Ten adult listeners (4 male, 6 female) participated in the experiment (age: $M = 23.0$, $SD = 2.3$). All subjects were classified as nonmusicians as assessed by a music history questionnaire (Wong and Perrachione, 2007) having no more than 3 years of formal musical training on any combination of instruments ($M = 0.9$, $SD = 1.08$). In addition, none of the participants possessed absolute (i.e., perfect) pitch nor had ever had musical ear training of any kind. All exhibited normal hearing sensitivity (i.e., better than 15 dB HL in both ears) at audiometric frequencies between 500 and 4000 Hz, and none reported any previous history of neurological or psychiatric illnesses. Each gave informed consent in compliance with a protocol approved by the Institutional Review Board of Purdue University.

Stimuli. A set of nine musical dyads (i.e., two-note musical intervals) were constructed to match those found in similar studies on consonance and dissonance (Kameoka and Kuriyagawa, 1969b, a). Individual notes were synthesized using a tone-complex consisting of 6 harmonics (equal amplitude, added in sine phase) whose fundamental frequency corresponded to a different pitch along the Western musical scale (Table 1; see also supplemental material, available at www.jneurosci.org). For every dyad, the lower of the two pitches was fixed with a fundamental frequency (f_0) of 220 Hz (A3 on the Western music scale), while the upper pitch was varied to produce different musical interval relationships within the range of an octave. Six consonant intervals (unison: ratio 1:1, major third: 5:4, perfect fourth: 4:3, perfect fifth: 3:2, major sixth: 5:3, octave: 2:1) and three dissonant intervals (minor second: 16:15, tritone: 45:32, major seventh: 15:8) were used in the experiment. Stimulus waveforms were 200 ms in duration including a 10 ms $\cos^2$ ramp applied at both the onset and offset to reduce both spectral splatter in the stimuli and onset components in the responses.

Behavioral judgments of consonance and dissonance. Subjective ratings of consonance and dissonance were measured using a paired comparison paradigm. Interval dyads were presented dichotically (one note per ear) to each participant at an intensity of ~70 dB SPL through circumaural headphones (Sennheiser HD 580). In each trial, listeners heard two musical intervals (one after the other) and were asked to select the interval they thought sounded more consonant (i.e., pleasant, beautiful, euphonious) (Plomp and Levelt, 1965) via a mouse click in a custom GUI coded in MATLAB 7.5 (The MathWorks). Participants were allowed to replay each interval within a given pair if they required more than one hearing before making a selection. The order of the two intervals in a pair combination was randomly assigned and all possible pairings were presented to each subject. In total, each individual heard 36 interval pairs (9 intervals, choose 2) such that each musical interval was contrasted with every other interval. A consonance rating for each dyad was then computed by counting the number of times it was selected relative to the total number of possible comparisons.

FFR recording protocol. Following the behavioral experiment, participants reclined comfortably in an acoustically and electrically shielded booth to facilitate recording of brainstem FFRs. They were instructed to relax and refrain from extraneous body movement to minimize myogenic artifacts. Subjects were allowed to sleep through the duration of the FFR experiment. FFRs were recorded from each participant in response to dichotic presentation of the nine stimuli (i.e., a single note was presented to each ear) at an intensity of 77 dB SPL through two magnetically shielded insert earphones (Etymotic ER-3A) (see supplemental Fig. 1, available at www.jneurosci.org as supplemental material). Dichotic presentation was used to minimize peripheral processing and isolate responses from a central (e.g., brainstem) pitch mechanism (Houtsma and Goldstein, 1972; Houtsma, 1979). Separating the notes between ears also ensured that distortion product components caused by cochlear nonlinearities would not obscure FFR responses (cf. Elsisy and Krishnan, 2008; Lee et al., 2009). Each stimulus was presented using single polarity and a repetition rate of 3.45/s. The presentation order of the stimuli was randomized both within and across participants. Control of the experimental protocol was accomplished by a signal generation and data acquisition system (Tucker-Davis Technologies, System III).

FFRs were recorded differentially between a noninverting (+) electrode placed on the midline of the forehead at the hairline (Fz) and an inverting (−) electrode placed on the seventh cervical vertebra (C7). Another electrode placed on the midforehead (Fpz) served as the common ground. All interelectrode impedances were maintained at or below 1 kΩ. The EEG inputs were amplified by 200,000 and bandpass filtered from 70 to 5000 Hz (6 dB/octave roll-off, RC response characteristics).
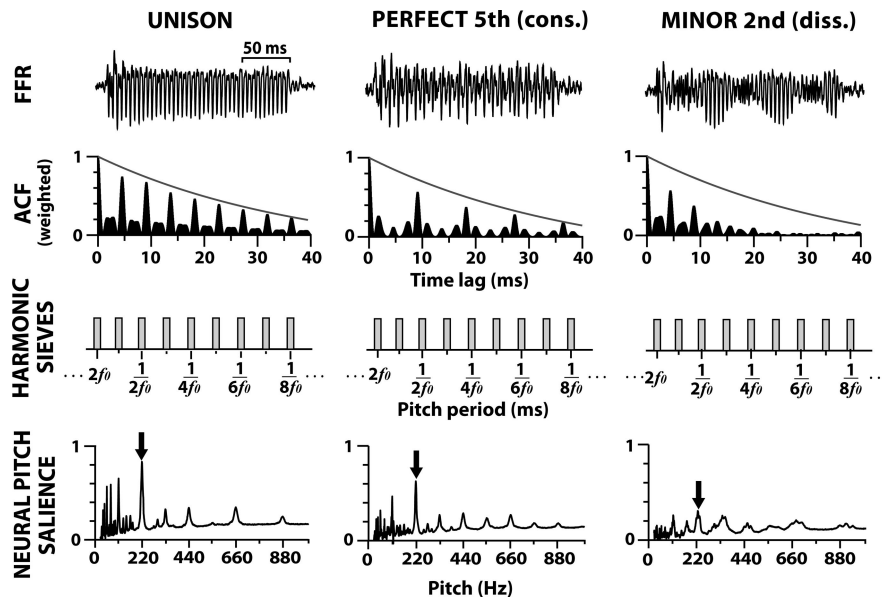
Figure 1. Procedure for computing neural pitch salience from FFR responses to musical intervals [unison, perfect fifth (consonant), and minor second (dissonant) shown here]. Dichotic presentation of a musical dyad elicits the scalp-recorded FFR response (top row). From each FFR waveform, the autocorrelation function (ACF) is computed and time weighted with a decaying exponential (solid gray line) to calculate the behaviorally relevant periodicities present in the response (second row). Each ACF is then passed through a series of harmonic interval pitch sieves consisting of "windows" centered at f_0 and its integer harmonics (third row). Each sieve template represents a single pitch and the magnitude of the output of each individual sieve represents a measure of neural pitch salience at that pitch. Analyzing the outputs across all possible pitches (25–1000 Hz) results in a running pitch salience for the stimulus (fourth row). As the arrows indicate, the magnitude of pitch salience for a consonant musical interval is more robust than that of a dissonant musical interval (e.g., compare perfect fifth to minor second). Yet, neither interval produces stronger neural pitch salience than the unison.

Sweeps containing activity exceeding ± 30 μV were rejected as artifacts. In total, each response waveform represents the average of 3000 artifact free trials over a 230 ms acquisition window.

FFR data analysis. Since the FFR reflects phase-locked activity in a population of neural elements, we adopted a temporal analysis scheme in which we examined the periodicity information contained in the aggregate distribution of neural activity (Langner, 1983; Rhode, 1995; Cariani and Delgutte, 1996b, a). We first computed the autocorrelation function (ACF) of each FFR. The ACF is equivalent to an all-order interspike interval histogram (ISIH) and represents the dominant pitch periodicities present in the neural response (Cariani and Delgutte, 1996a; Krishnan et al., 2004). Each ACF was then weighted with a decaying exponential ($\tau = 10$ ms) to give greater precedence to shorter pitch intervals (Cedolin and Delgutte, 2005) and account for the lower limit of musical pitch (~30 Hz) (Pressnitzer et al., 2001).

To obtain an estimate of the neural pitch salience, we then analyzed each FFR's weighted ACF using "period template" analysis. A series of dense harmonic interval sieves was applied to each weighted ACF (i.e., ISIH) to quantify the neural activity at a given pitch period and its multiples (Cedolin and Delgutte, 2005; Larsen et al., 2008). Each sieve template (representing a single pitch) was composed of 500 μs wide bins situated at the fundamental pitch period ($1/f_0$) and its integer multiples. All sieve templates with f_0s between 25 and 1000 Hz (2 Hz steps) were used in the analysis. The salience for a given pitch was estimated by dividing the mean density of neural activity falling within the sieve bins by the mean density of activity in the whole distribution. ACF activity falling within sieve "windows" adds to the total pitch salience while information falling outside the "windows" reduces the total pitch salience (Cedolin and Delgutte, 2005). By compounding the output of all sieves as a function of f_0 we examine the relative strength of all possible pitches present in the FFR which may be associated with different perceived pitches. The pitch (f_0) yielding maximum salience was taken as the "heard pitch" (Parncutt, 1989) which in every case was 220 Hz. To this end, we measured the pitch salience magnitude at 220 Hz for each FFR in

response to each musical interval stimulus. This procedure allowed us to directly compare the pitch salience between different musical intervals and quantitatively measure their differential encoding. The various steps in the harmonic sieve analysis procedure are illustrated in Figure 1. In addition, we computed narrowband FFTs (resolution = 5 Hz, via zero-padding) to evaluate the spectral composition of each FFR.

Statistical analysis. Neural pitch salience derived from the FFRs was analyzed using a mixed-model ANOVA (SAS) with subjects as a random factor and musical interval as the within-subject, fixed effect factor (nine levels; Un, m2, M3, P4, TT, P5, M6, M7, Oct) to assess whether pitch encoding differed between musical intervals. By examining the robustness of neural pitch salience across musical intervals, we determine whether subcortical pitch processing shows differential sensitivity in encoding the basic pitch relationships found in music.

Results

FFRs preserve complex spectra of multiple musical pitches

Grand-average FFRs and their corresponding frequency spectra are shown for a subset of the stimuli in Figure 2, *A* and *B*, respectively. As evident by the spectral magnitudes, FFRs to consonant musical intervals produce more robust neural responses than dissonant intervals. Unlike consonant intervals, temporal waveforms of dissonant intervals are characterized by a "beating" effect (evident by their modulated temporal envelopes) caused by the difference tone between their nearby fundamental frequencies (e.g., compare unison to minor second). In the case of the minor second, beating produces amplitude minima every $1/\Delta f = 67$ ms, where Δf is the difference frequency of the two pitches (i.e., 235–220 Hz = 15 Hz). Importantly, though dyads were presented dichotically (one note to each ear), FFRs preserve the complex spectra of both notes in a single response (compare response spectra, filled areas, to stimulus spectra, harmonic locations denoted by dots). While some frequency components of the input stimuli exceed 2500 Hz, no spectral information was present in the FFRs above ~1000 Hz, consistent with the upper limit of phase locking in the IC (Liu et al., 2006).

Behavioral ratings of consonance and dissonance

Individual behavioral consonance ratings for each interval were obtained by computing the proportion of times a given interval was selected by a participant out of the 36 total pairwise comparisons. Mean behavioral consonance ratings for the nine musical intervals are displayed in Figure 3*A*. Subjects generally selected consonant intervals more frequently than dissonant intervals, suggesting that the former was judged to be more pleasant sounding than the latter. In particular, perfect consonant intervals (unison, perfect fourth, fifth, and octave) were rated higher than imperfect consonant intervals (major third and major sixth). However, regardless of quality, consonance intervals (solid bars) were judged more pleasant than dissonant intervals (hatched bars; minor second, tritone, and major seventh). The ordering of consonance observed here is consistent with previous reports of

musical interval ratings (Malmberg, 1918; Plomp and Levelt, 1965; Kameoka and Kuriyagawa, 1969b; Krumhansl, 1990; Itoh et al., 2003a; Schwartz et al., 2003). Our results are also consistent with the predicted hierarchical ordering of intervals obtained from a mathematical model of sensory consonance and dissonance (solid curve) (Sethares, 1993). In this model, interacting harmonic components between the two tones produce either constructive or destructive interference. Higher degrees of consonance are represented by local maxima, dissonance by local minima.

Somewhat surprising is the fact that listeners on average, rated the octave and perfect fifth higher than the unison, which by definition, is the most consonant interval attainable (i.e., two notes of the exact same pitch) (see Fig. 3A). However, this pattern is often observed in nonmusicians when compared with their musician counterparts (van de Geer et al., 1962). As such, our participants have no prior exposure to the rules of music theory and therefore, do not have internal, preconceived labels or categorizations for what is "consonant" or "dissonant" (van de Geer et al., 1962; Burns and Ward, 1978). When asked to compare between the unison and other intervals, it is likely that our listeners chose the latter because they heard two distinct pitches. In other words, because the unison does not sound like an "interval" per se, individuals probably avoided selecting it in favor of the alternative, more interesting-sounding interval in the pair they were comparing.
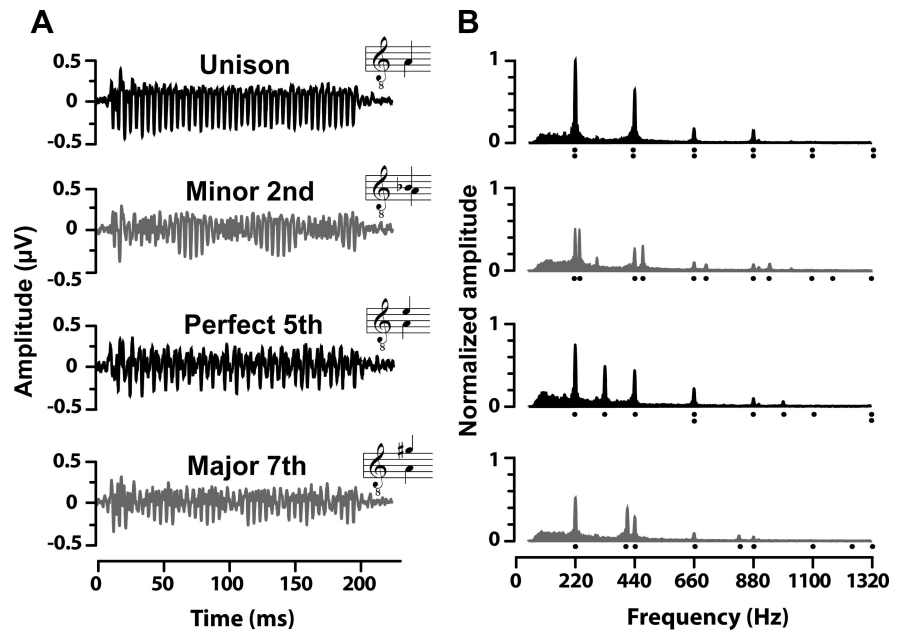

Figure 2. Grand-average FFR waveforms (**A**) and their corresponding frequency spectra (**B**) elicited from the dichotic presentation of four representative musical intervals. Consonant intervals are shown in black, dissonant intervals in gray. **A**, Time waveforms reveal clearer periodicity and more robust amplitudes for consonant intervals than dissonant intervals. In addition, dissonant dyads (e.g., minor second and major seventh) show significant interaction of frequency components as evident from the modulated nature of their waveforms. Insets show the musical notation for the input stimulus. **B**, Frequency spectra reveal that FFRs faithfully preserve the harmonic constituents of both musical notes even though they were presented separately between the two ears (compare response spectrum, filled area, to stimulus spectrum, harmonic locations denoted by dots). Consonant intervals have higher spectral magnitudes across harmonics than dissonant intervals. All amplitudes are normalized relative to the unison.

Brainstem pitch salience reveals differential encoding of musical intervals
Mean neural pitch salience derived from individual FFRs are shown for each of the nine musical interval stimuli in Figure 3B. Results of the omnibus ANOVA showed a significant effect of musical interval on neural pitch salience ($F_{(8,72)} = 17.47$, $p < 0.0001$). Post hoc Tukey–Kramer adjustments ($\alpha = 0.05$) were used to evaluate all pairwise multiple comparisons. Results of these tests revealed that the unison elicited significantly larger neural pitch salience than all other musical intervals except the octave, its closest relative. The minor second, a highly dissonant interval, produced significantly lower pitch salience than the perfect fourth, perfect fifth, tritone, major sixth, and octave. As expected, the two most dissonant intervals (minor second and major seventh) did not differ in pitch salience. The perfect fifth yielded higher neural pitch salience than the major third and major seventh. The octave, which was judged highly consonant in the behavioral task, yielded larger neural pitch salience than all of the dissonant intervals (minor second, tritone, major seventh) and two of the consonant intervals (major third and major sixth). However, the octave did not differ from the unison, perfect fourth, or perfect fifth. Interestingly, these four intervals were the same pitch combinations which also produced the highest behavioral consonance ratings. These intervals are also typically considered most structural, or "pure," in tonal music. An independent samples t test (two-tailed) was used to directly compare all con-

sonant musical intervals (unison, major third, perfect fourth, perfect fifth, major sixth, octave) to all dissonant musical intervals (minor second, tritone, major seventh). Results of this contrast revealed that, on the whole, consonant intervals produced significantly stronger neural pitch salience than dissonant intervals ($t_{(88)} = 4.28$, $p < 0.0001$).

Behavioral versus neural data
Figure 4 shows behavioral consonance ratings plotted against neural pitch salience for each musical interval. Neural and behavioral data showed a significant correlation (Pearson's $r = 0.81$, $p = 0.0041$), suggesting that subcortical processing can, in part, predict an individual's behavioral judgments of musical pitch relationships. Consonant musical intervals judged more pleasant by the listeners also yielded more robust neural pitch salience than dissonant intervals, as evident by the clustering of the two categories. Intervals deemed most consonant by music theory standards (e.g., unison and octave) are separated maximally in distance from those deemed highly dissonant (e.g., minor second and major seventh). Although musical training is known to enhance the subcortical processing of musically relevant pitch (Musacchia et al., 2007; Bidelman et al., 2009; Lee et al., 2009), we note that these relationships are represented even in nonmusicians.

Discussion
The results of this study relate to three main observations: (1) neural phase-locked activity in the brainstem appears to preserve information relevant to the perceptual attributes of consonance, dissonance, and the hierarchical relations of musical pitch; (2) the strength of aggregate neural activity within the brainstem appears to be correlated with the relative ordering of consonance found behaviorally;
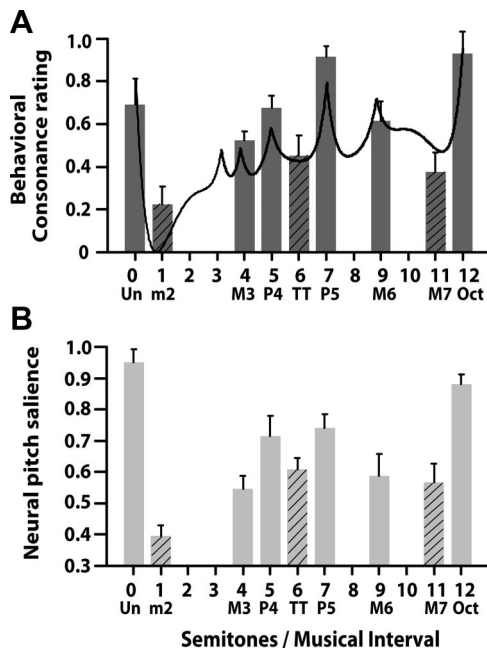
Figure 3. Perceptual consonant ratings of musical intervals and estimates of neural pitch salience derived from their respective FFRs. Solid bars, Consonant intervals; hatched bars, dissonant intervals. *A*, Mean behavioral consonance ratings for dichotic presentation of nine musical intervals. Dyads considered consonant according to music theory (solid bars) are preferred over those considered dissonant [minor second (m2), tritone (TT), major seventh (M7)]. For comparison, the solid line shows predictions from a mathematical model of consonance and dissonance (Sethares, 1993) in which local maxima denote higher degrees of consonance than minima, which denote dissonance. *B*, Mean neural pitch salience derived from FFR responses to dichotic musical intervals. Consonant intervals produce greater pitch salience than dissonant intervals. Even among intervals common to a single class (e.g., all consonant intervals) FFRs show differential encoding resulting in the hierarchical arrangement of pitch described by Western music theory. Error bars indicate one SEM.
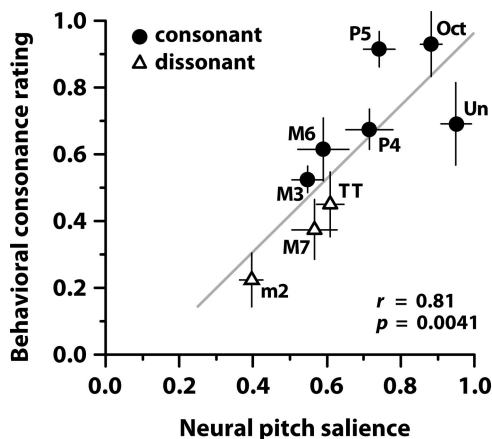

Figure 4. Neural pitch salience derived from FFRs versus behavioral consonance ratings. Consonant intervals elicit a larger neural pitch salience than dissonant intervals and are judged more pleasant by the listener. Note the systematic clustering of consonant and dissonant intervals and the maximal separation of the unison (most consonant interval) from the minor second (most dissonant interval). Error bars indicate one SEM in either the behavioral or neural dimension, respectively.

and (3) basic properties of musical pitch structure are encoded even in individuals without formal musical training.

Across individuals, consonant musical intervals were characterized by more robust responses which yielded stronger neural pitch salience than dissonant intervals, an intuitive finding not

observed in previous studies examining brainstem responses to musical intervals (cf. Lee et al., 2009). Perceptually, these same intervals were judged to be more pleasant sounding by our listeners. It is important to point out that the nature of neural activity we observe is graded. The same can be said about our behavioral data. That is, musical pitch relationships are not encoded in a strict binary manner (i.e., consonant vs dissonant) but rather, are processed differentially based on their degree of perceptual consonance. In the present data, this is evident by the fact that even within a given class (e.g., all consonant dyads), intervals elicit graded levels of pitch salience and lead to different subjective ratings (e.g., compare consonant perfect fifth and consonant major third) (Fig. 3*B*). Such a differential encoding scheme may account for the hierarchical ranking of musical intervals measured here and in past psychophysical studies (Krumhansl, 1990; Schwartz et al., 2003).

Psychophysical basis of consonance and dissonance
It was recognized as early as Pythagoras that pleasant (i.e., consonant)-sounding musical intervals were produced when the frequencies of two vibrating entities formed simple integer ratios (e.g., 3:2 corresponds to the perfect fifth, 2:1 the octave) (see Table 1), whereas complex ratios produced "harsh"- or "rough"-sounding tones (e.g., 16:15, dissonant minor second). Helmholtz (1877) later postulated that dissonance arose as a result of interference between frequency components within two or more tones. That is, when two pitches are spaced close together in frequency, their harmonics interfere and create the perception of "roughness" or "beating," as typically described for dissonant tones. Consonance presumably occurs in the absence of beating, when harmonics are spaced sufficiently far apart so as to not interact. Empirical studies have suggested this phenomenon is related to cochlear mechanics and the so-called critical band hypothesis (Plomp and Levelt, 1965). This theory infers that the overall consonance of a musical interval depends on the total interaction of frequency components within single auditory filters. Pitches of consonant dyads have fewer partials that pass through the same critical band and therefore yield more pleasant percepts than dissonant intervals whose partials compete within individual channels. However, this psychoacoustic model fails to explain the perception of consonance and dissonance when tones are presented dichotically (Houtsma and Goldstein, 1972), as in the present study. In these cases, pitch percepts must be computed centrally by deriving information from the combined neural signals relayed from both cochleae.

Physiological basis of consonance and dissonance
Animal studies provide physiological evidence for a critical band mechanism operating in the brainstem. Tonotopically organized frequency lamina within the cat IC exhibit constant frequency ratios between adjacent layers (Schreiner and Langner, 1997). This anatomy, coupled with extensive lateral inhibition, provides evidence for a midbrain critical band that may function in a similar manner to that described psychophysically (Plomp and Levelt, 1965; Braun, 1999). One of the possible perceptual consequences for this type of architecture may be a sensitivity to consonant and dissonant pitch relationships (Braun, 1999; Schwartz et al., 2003). Indeed, our FFR results suggest that the IC may act as a driving input to the so-called "central pitch processor" proposed by Houtsma and Goldstein (1972) in their studies on dichotic musical pitch.

Our results offer evidence that musical pitch percepts may ultimately arise from temporally based processing schemes

(Langner, 1997). The high correspondence between neural pitch salience derived from brainstem FFRs and behavioral consonance ratings suggests that the perceptual pleasantness of musical units may originate from the temporal distribution of firing patterns within subcortical structures (Tramo et al., 2001). We found that consonant pitch intervals generate more robust and synchronous phase locking than dissonant pitch intervals, consistent with previous results from cat auditory nerve (Tramo et al., 2001) and IC (McKinney et al., 2001). For consonant dyads, interspike intervals within the population activity occur at precise, harmonically related pitch periods thereby producing maximal periodicity in their neural representation. Dissonant intervals on the other hand produce multiple, more irregular neural periodicities. Pitch-related mechanisms operating in the brainstem likely use simple periodic (cf. consonant) information more effectively than aperiodic (cf. dissonant) information (Rhode, 1995; Langner, 1997; Ebeling, 2008), as the former is likely to be more compatible with pitch extraction templates and provides a more robust, unambiguous cue for pitch (McDermott and Oxenham, 2008). In a sense, then, dissonance may challenge the auditory system in ways that simple consonance does not. Indeed, consonant intervals may ultimately reduce computational load and/or require fewer brain resources to process than their dissonant counterparts due to the more coherent, synchronous neural activity they evoke (Burns, 1999, p. 243).

To date, very few studies have investigated the role of subcortical sensory-level processing in forming the perceptual attributes related to musical pitch structure [for exceptions, see Tramo et al. (2001) and Lee et al. (2009)]. There is however, overwhelming evidence to suggest that cortical integrity is necessary for maintaining perceptual faculties related to the processing of musical pitch (Johnsrude et al., 2000; Ayotte et al., 2002; Peretz et al., 2009). Tonal scale structures are represented topographically in cerebral cortex (Janata et al., 2002), and cortical potentials known to index attentional and associative processes show differential sensitivity to specific pitch relationships that occur in music (Itoh et al., 2003b; Krohn et al., 2007). In addition, studies using the mismatched negativity (MMN) and similar early-latency ERPs have demonstrated that interval-dependent effects occur automatically at a preattentive stage of processing (Brattico et al., 2006; Bergelson and Idsardi, 2009). Together, these studies offer compelling evidence that the hierarchical status of musical pitch is maintained at a cortical level. Our results now extend this framework to a subcortical level. Brain networks engaged during music likely involve a series of computations applied to the neural representation at different stages of processing (Hickok and Poeppel, 2004). We argue that higher-level abstract representations of musical pitch hierarchy previously examined only cortically are grounded in sensory features that emerge very early along the auditory pathway. While the formation of a musical interval percept may ultimately lie with cortical mechanisms, subcortical structures are clearly feeding this architecture with relevant information. To our knowledge, this is the first time that the hierarchical representation of musical pitch has been demonstrated at a subcortical level in humans (cf. Tramo et al., 2001).

Is there a neurobiological basis for the structure of musical pitch?

There are notable commonalities (i.e., universals) among many of the music systems of the world including the division of the octave into specific scale steps and the use of a stable reference pitch to establish key structure. In fact, it has been argued that culturally specific music is simply an elaboration of only a few

universal traits (Carterette and Kendall, 1999), one of which is the preference for consonance (Fritz et al., 2009). Our results imply that the perceptual attributes related to such preferences may be a byproduct of innate sensory-level processing. Though we only measured adult nonmusicians, behavioral studies on infants have drawn similar conclusions showing that even months into life, newborns prefer listening to consonant rather than dissonant musical sequences (Trainor et al., 2002) and tonal versus atonal melodies (Trehub et al., 1990). Given that these effects are observed in the absence of long-term enculturation, exposure, or music training, it is conceivable that the perception of pitch structure develops from domain-general processing governed by the fundamental capabilities of the auditory system (Tramo et al., 2001; McDermott and Hauser, 2005; Zatorre and McGill, 2005; Trehub and Hannon, 2006; Trainor, 2008). It is interesting to note that intervals deemed more pleasant sounding by listeners are also more prevalent in tonal composition (Budge, 1943; Vos and Troost, 1989; Huron, 1991). A neurobiological predisposition for simpler, consonant intervals may be one reason why such pitch combinations have been favored by composers and listeners for centuries (Burns, 1999). Indeed, the very arrangement of musical notes into a hierarchical structure may be a consequence of the fact that certain pitch combinations strike a deep chord with the architecture of the nervous system (McDermott, 2008).

Conclusion

By examining the subcortical response to musical intervals we found that consonance, dissonance, and the hierarchical ordering of musical pitch are automatically encoded by preattentive, sensory-level processing. Brainstem responses are well correlated with the ordering of consonance obtained behaviorally suggesting that a listener's judgment of pleasant- or unpleasant-sounding music may be rooted in low-level sensory processing. Though music training is known to tune cortical and subcortical circuitry, the fundamental attributes of musical pitch we examined here are encoded even in nonmusicians. It is possible that the choice of intervals used in compositional practice may have originated based on the fundamental processing and constraints of the auditory system.

References

Ayotte J, Peretz I, Hyde K (2002) Congenital amusia: a group study of adults afflicted with a music-specific disorder. Brain 125:238–251.

Bergelson E, Idsardi WJ (2009) A neurophysiological study into the foundations of tonal harmony. Neuroreport 20:239–244.

Besson M, Faita F (1995) An event-related potential (ERP) study of musical expectancy: comparison of musicians with nonmusicians. J Exp Psychol Hum Percept Perform 21:1278–1296.

Bidelman GM, Gandour JT, Krishnan A (2009) Cross-domain effects of music and language experience on the representation of pitch in the human brain. J Cogn Neurosci, in press.

Brattico E, Tervaniemi M, Näätänen R, Peretz I (2006) Musical scale properties are automatically processed in the human auditory cortex. Brain Res 1117:162–174.

Braun M (1999) Auditory midbrain laminar structure appears adapted to f0 extraction: further evidence and implications of the double critical bandwidth. Hear Res 129:71–82.

Budge H (1943) A study of chord frequencies. New York: Bureau of Publications, Teachers College, Columbia University.

Burns EM (1999) Intervals, scales, and tuning. In: The psychology of music, Ed 2 (Deutsch D, ed), pp 215–264. San Diego: Academic.

Burns EM, Ward WD (1978) Categorical perception—phenomenon or epiphenomenon: evidence from experiments in the perception of melodic musical intervals. J Acoust Soc Am 63:456–468.

Cariani PA, Delgutte B (1996a) Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. J Neurophysiol 76:1698–1716.

Cariani PA, Delgutte B (1996b) Neural correlates of the pitch of complex tones. II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region for pitch. J Neurophysiol 76:1717–1734.

Carterette EC, Kendall RA (1999) Comparative music perception and cognition. In: The psychology of music, Ed 2 (Deutsch D, ed), pp 725–791. San Diego: Academic.

Cedolin L, Delgutte B (2005) Pitch of complex tones: rate-place and inter-spike interval representations in the auditory nerve. J Neurophysiol 94:347–362.

Ebeling M (2008) Neuronal periodicity detection as a basis for the perception of consonance: a mathematical model of tonal fusion. J Acoust Soc Am 124:2320–2329.

Elsisy H, Krishnan A (2008) Comparison of the acoustic and neural distortion product at 2f1–f2 in normal-hearing adults. Int J Audiol 47:431–438.

Fishman YI, Volkov IO, Noh MD, Garell PC, Bakken H, Arezzo JC, Howard MA, Steinschneider M (2001) Consonance and dissonance of musical chords: neural correlates in auditory cortex of monkeys and humans. J Neurophysiol 86:2761–2788.

Fritz T, Jentschke S, Gosselin N, Sammler D, Peretz I, Turner R, Friederici AD, Koelsch S (2009) Universal recognition of three basic emotions in music. Curr Biol 19:573–576.

Galbraith GC, Threadgill MR, Hemsley J, Salour K, Songdej N, Ton J, Cheung L (2000) Putative measure of peripheral and brainstem frequency-following in humans. Neurosci Lett 292:123–127.

Hannon EE, Trainor LJ (2007) Music acquisition: effects of enculturation and formal training on development. Trends Cogn Sci 11:466–472.

Helmholtz H (1877) On the sensations of tone. Reprint (Ellis AJ, translator). New York: Dover, 1954.

Hickok G, Poeppel D (2004) Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition 92:67–99.

Houtsma AJ (1979) Musical pitch of two-tone complexes and predictions by modern pitch theories. J Acoust Soc Am 66:87–99.

Houtsma AJ, Goldstein JL (1972) The central origin of the pitch of complex tones: evidence from musical interval recognition. J Acoust Soc Am 51:520–529.

Huron D (1991) Tonal consonance versus tonal fusion in polyphonic sonorities. Music Percept 9:135–154.

Itoh K, Miyazaki K, Nakada T (2003a) Ear advantage and consonance of dichotic pitch intervals in absolute-pitch possessors. Brain Cogn 53:464–471.

Itoh K, Suwazono S, Nakada T (2003b) Cortical processing of musical consonance: an evoked potential study. Neuroreport 14:2303–2306.

Janata P, Birk JL, Van Horn JD, Leman M, Tillmann B, Bharucha JJ (2002) The cortical topography of tonal structures underlying Western music. Science 298:2167–2170.

Johnsrude IS, Penhune VB, Zatorre RJ (2000) Functional specificity in the right human auditory cortex for perceiving pitch direction. Brain 123:155–163.

Kameoka A, Kuriyagawa M (1969a) Consonance theory part I: consonance of dyads. J Acoust Soc Am 45:1451–1459.

Kameoka A, Kuriyagawa M (1969b) Consonance theory part II: consonance of complex tones and its calculation method. J Acoust Soc Am 45:1460–1469.

Koelsch S, Gunter T, Friederici AD, Schröger E (2000) Brain indices of music processing: "nonmusicians" are musical. J Cogn Neurosci 12:520–541.

Krishnan A, Gandour JT (2009) The role of the auditory brainstem in processing linguistically-relevant pitch patterns. Brain Lang 110:135–148.

Krishnan A, Xu Y, Gandour JT, Cariani PA (2004) Human frequency-following response: representation of pitch contours in Chinese tones. Hear Res 189:1–12.

Krohn KI, Brattico E, Välimäki V, Tervaniemi M (2007) Neural representations of the hierarchical scale pitch structure. Music Percept 24:281–296.

Krumhansl CL (1990) Cognitive foundations of musical pitch. New York: Oxford UP.

Langner G (1983) Evidence for neuronal periodicity detection in the auditory system of the guinea fowl: implications for pitch analysis in the time domain. Exp Brain Res 52:333–355.

Langner G (1997) Neural processing and representation of periodicity pitch. Acta Otolaryngol Suppl 532:68–76.

Larsen E, Cedolin L, Delgutte B (2008) Pitch representations in the auditory nerve: two concurrent complex tones. J Neurophysiol 100:1301–1319.

Lee KM, Skoe E, Kraus N, Ashley R (2009) Selective subcortical enhancement of musical intervals in musicians. J Neurosci 29:5832–5840.

Liu LF, Palmer AR, Wallace MN (2006) Phase-locked responses to pure tones in the inferior colliculus. J Neurophysiol 95:1926–1935.

Malmberg CF (1918) The perception of consonance and dissonance. Psychol Monogr 25:93–133.

McDermott J (2008) The evolution of music. Nature 453:287–288.

McDermott J, Hauser MD (2005) The origins of music: innateness, uniqueness, and evolution. Music Percept 23:29–59.

McDermott JH, Oxenham AJ (2008) Music perception, pitch, and the auditory system. Curr Opin Neurobiol 18:452–463.

McKinney MF, Tramo MJ, Delgutte B (2001) Neural correlates of the dissonance of musical intervals in the inferior colliculus. In: Physiological and psychophysical bases of auditory function (Breebaart DJ, Houtsma AJM, Kohlrausch A, Prijs VF, Schoonhoven R, eds), pp 83–89. Maastricht, The Netherlands: Shaker.

Minati L, Rosazza C, D'Incerti L, Pietrocini E, Valentini L, Scaioli V, Loveday C, Bruzzone MG (2008) FMRI/ERP of musical syntax: comparison of melodies and unstructured note sequences. Neuroreport 19:1381–1385.

Musacchia G, Sams M, Skoe E, Kraus N (2007) Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. Proc Natl Acad Sci U S A 104:15894–15898.

Parncutt R (1989) Harmony: a psychoacoustical approach. Berlin: Springer.

Peretz I, Brattico E, Järvenpää M, Tervaniemi M (2009) The amusic brain: in tune, out of key, and unaware. Brain 132:1277–1286.

Plomp R, Levelt WJ (1965) Tonal consonance and critical bandwidth. J Acoust Soc Am 38:548–560.

Pressnitzer D, Patterson RD, Krumbholz K (2001) The lower limit of melodic pitch. J Acoust Soc Am 109:2074–2084.

Rameau JP (1722) Treatise on harmony. Reprint. New York: Dover, 1971.

Rhode WS (1995) Interspike intervals as a correlate of periodicity pitch in cat cochlear nucleus. J Acoust Soc Am 97:2414–2429.

Schreiner CE, Langner G (1997) Laminar fine structure of frequency organization in auditory midbrain. Nature 388:383–386.

Schwartz DA, Howe CQ, Purves D (2003) The statistical structure of human speech sounds predicts musical universals. J Neurosci 23:7160–7168.

Sethares WA (1993) Local consonance and the relationship between timbre and scale. J Acoust Soc Am 94:1218–1228.

Smith JC, Marsh JT, Brown WS (1975) Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. Electroencephalogr Clin Neurophysiol 39:465–472.

Trainor L (2008) Science and music: the neural roots of music. Nature 453:598–599.

Trainor LJ, Tsang CD, Cheung VHW (2002) Preference for sensory consonance in 2- and 4-month-old infants. Music Percept 20:187–194.

Tramo MJ, Cariani PA, Delgutte B, Braida LD (2001) Neurobiological foundations for the theory of harmony in western tonal music. Ann N Y Acad Sci 930:92–116.

Trehub SE, Hannon EE (2006) Infant music perception: domain-general or domain-specific mechanisms? Cognition 100:73–99.

Trehub SE, Thorpe LA, Trainor LJ (1990) Infants' perception of good and bad melodies. Psychomusicology 9:5–19.

van de Geer JP, Levelt WJM, Plomp R (1962) The connotation of musical consonance. Acta Psychol 20:308–319.

Vos PG, Troost JM (1989) Ascending and descending melodic intervals: statistical findings and their perceptual relevance. Music Percept 6:383–396.

Wong PC, Perrachione TK (2007) Learning pitch patterns in lexical identification by native English-speaking adults. Appl Psycholinguist 28:565–585.

Wong PC, Skoe E, Russo NM, Dees T, Kraus N (2007) Musical experience shapes human brainstem encoding of linguistic pitch patterns. Nat Neurosci 10:420–422.

Zatorre R, McGill J (2005) Music, the food of neuroscience? Nature 434:312–315.