

# Major and minor music compared to excited and subdued speech

Daniel L. Bowling, Kamraan Gill, Jonathan D. Choi, Joseph Prinz, and Dale Purves  
*Department of Neurobiology and Center for Cognitive Neuroscience, Duke University Durham, North Carolina 27708*

(Received 3 June 2009; revised 15 October 2009; accepted 16 October 2009)

The affective impact of music arises from a variety of factors, including intensity, tempo, rhythm, and tonal relationships. The emotional coloring evoked by intensity, tempo, and rhythm appears to arise from association with the characteristics of human behavior in the corresponding condition; however, how and why particular tonal relationships in music convey distinct emotional effects are not clear. The hypothesis examined here is that major and minor tone collections elicit different affective reactions because their spectra are similar to the spectra of voiced speech uttered in different emotional states. To evaluate this possibility the spectra of the intervals that distinguish major and minor music were compared to the spectra of voiced segments in excited and subdued speech using fundamental frequency and frequency ratios as measures. Consistent with the hypothesis, the spectra of major intervals are more similar to spectra found in excited speech, whereas the spectra of particular minor intervals are more similar to the spectra of subdued speech. These results suggest that the characteristic affective impact of major and minor tone collections arises from associations routinely made between particular musical intervals and voiced speech.

© 2010 Acoustical Society of America. [DOI: 10.1121/1.3268504]

PACS number(s): 43.75.Cd, 43.75.Bc, 43.71.Es [DD]

Pages: 491–503

## I. INTRODUCTION

The affective impact of music depends on many factors including, but not limited to, intensity, tempo, rhythm, and the tonal intervals used. For most of these factors the way emotion is conveyed seems intuitively clear. If, for instance, a composer wants to imbue a composition with excitement, the intensity tends to be forte, the tempo fast, and the rhythm syncopated; conversely, if a more subdued effect is desired, the intensity is typically piano, the tempo slower, and the rhythm more balanced (Cohen, 1971; Bernstein, 1976; Juslin and Laukka, 2003). These effects on the listener presumably occur because in each case the characteristics of the music accord with the ways that the corresponding emotional state is expressed in human behavior. The reason for the emotional effect of the tonal intervals used in music, however, is not clear.

Much music worldwide employs subsets of the chromatic scale, which divides each octave into 12 intervals defined by the frequency ratios listed in Table IA (Nettl, 1956; Randel, 1986; Carterette and Kendall, 1999; Burkholder *et al.*, 2005). Among the most commonly used subsets are the diatonic scales in Table IB (Pierce, 1962; Bernstein, 1976; Randel, 1986; Burns, 1999; Burkholder *et al.*, 2005). In recent centuries, the most popular diatonic scales have been the Ionian and the Aeolian, usually referred to simply as the major and the minor scale, respectively (Aldwell and Schachter, 2003) (Fig. 1). Other things being equal (e.g., intensity, tempo, and rhythm), music using the intervals of the major scale tends to be perceived as relatively excited, happy, bright, or martial, whereas music using minor scale intervals tends to be perceived as more subdued, sad, dark, or wistful (Zarlino, 1571; Hevner, 1935; Cooke, 1959; Crow-

der, 1984; Krumhansl, 1990; Gregory and Varney, 1996; Peretz *et al.*, 1998; Burkholder *et al.*, 2005). There has been, however, no agreement about how and why these scales and the intervals that differentiate them elicit distinct emotional effects (Heinlein, 1928; Hevner, 1935; Crowder, 1984; Carterette and Kendall, 1989; Schellenberg *et al.*, 2000; Gabriellsson and Lindstöm, 2001). Based on the apparent role of behavioral mimicry in the emotional impact produced by other aspects of music, we here examine the hypothesis that the same framework pertains to the affective impact of the different tone collections used in melodies.

To evaluate the merits of this hypothesis, we compared the spectra of the intervals that, based on an empirical analysis, specifically distinguish major and minor music with the spectra of voiced speech uttered in different emotional states. There are several reasons for taking this approach. First, voiced speech sounds are harmonic and many of the ratios between the overtones in any harmonic series correspond to musical ratios (Helmholtz 1885; Bernstein, 1976; Rossing 1990; Crystal, 1997; Stevens, 1998; Johnston, 2002). Second, most of the frequency ratios of the chromatic scale are statistically apparent in voiced speech spectra in a variety of languages (Schwartz *et al.*, 2003). Third, we routinely extract biologically important information about the emotional state of a speaker from the quality of their voice (Johnstone and Scherer, 2000; Scherer *et al.*, 2001; Juslin and Laukka, 2003; Thompson and Balkwill, 2006). Fourth, the physiological differences between excited and subdued affective states alter the spectral content of voiced speech (Spencer, 1857; Juslin and Laukka, 2003; Scherer, 2003). Fifth, with few exceptions, the only tonal sounds in nature are the vocalizations of animals, the most important of these being conspecific (Schwartz *et al.*, 2003). Sixth, a number of non-musical phe-

TABLE I. Western musical scales (also called modes). (A) The 12 intervals of the chromatic scale showing the abbreviations used, the corresponding number of semitones, and the ratio of the fundamental frequency of the upper tone to the fundamental frequency of the lower tone in just intonation tuning. (B) The seven diatonic scales/modes. As a result of their relative popularity, the Ionian and the Aeolian modes are typically referred to today as the major and minor scales, respectively. Although the Ionian and Aeolian modes and the scales they represent have been preeminent in Western music since the late 16th century, some of the other scales/modes continue to be used today. For example, the Dorian mode is used in plainchant and some folk music, the Phrygian mode is used in flamenco music, and the Mixolydian mode is used in some jazz. The Locrian and Lydian are rarely used because the dissonant tritone takes the place of the fifth and fourth scale degrees, respectively.

(A) Chromatic scale			(B) Diatonic scales						
Interval Name	Semitones	Frequency ratio	“MAJOR” Ionian	Dorian	Phrygian	Lydian	Mixolydian	“MINOR” Aeolian	Locrian
Unison (Uni)	0	1:1	M2	M2	m2	M2	M2	M2	m2
Minor second (m2)	1	16:15	M3	m3	m3	M3	M3	m3	m3
Major second (M2)	2	9:8	P4	P4	P4	tt	P4	P4	P4
Minor third (m3)	3	6:5	P5	P5	P5	P5	P5	P5	tt
Major third (M3)	4	5:4	M6	M6	m6	M6	M6	m6	m6
Perfect fourth (P4)	5	4:3	M7	m7	m7	M7	m7	m7	m7
Tritone (tt)	6	7:5	Oct	Oct	Oct	Oct	Oct	Oct	Oct
Perfect fifth (P5)	7	3:2							
Minor sixth (m6)	8	8:5							
Major sixth (M6)	9	5:3							
Minor seventh (m7)	10	9:5							
Major seventh (M7)	11	15:8							
Octave (Oct)	12	2:1							

nomena in pitch perception, including perception of the missing fundamental, the pitch shift of the residue, spectral dominance, and pitch strength can be rationalized in terms of spectral similarity to speech (Terhardt, 1974; Schwartz and Purves, 2004). And finally, as already mentioned, other aspects of music appear to convey emotion through mimicry of human behaviors that signify emotional state. It therefore makes sense to ask whether spectral differences that specifically distinguish major and minor melodies parallel spectral differences that distinguish excited and subdued speech.

## II. METHODS

### A. Overview

The intervals that distinguish major and minor music were determined from classical and folk melodies composed in major and minor keys. The notes in these melodies were

extracted and the intervals they represent were calculated and tallied by condition. Excited and subdued speech samples were obtained by recording single words and monologs spoken in either an excited or subdued manner. From these recordings only the voiced segments were extracted and analyzed. The spectra of the distinguishing musical intervals were then compared with the spectra of the voiced segments in excited and subdued speech according to fundamental frequency and frequency ratios.

### B. Acquisition and analysis of the musical databases

A database of classical Western melodies composed in major and minor keys over the past three centuries was compiled from the electronic counterpart of Barlow and Morgenstern’s Dictionary of Musical Themes (Barlow and Morgenstern, 1948). The online version (<http://www.multimedialibrary.com/Barlow>) includes 9825 monophonic

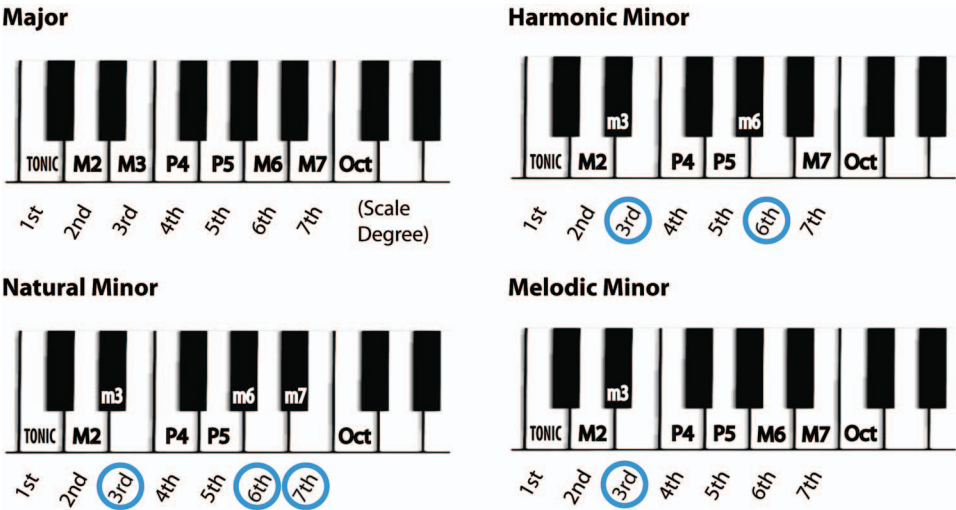
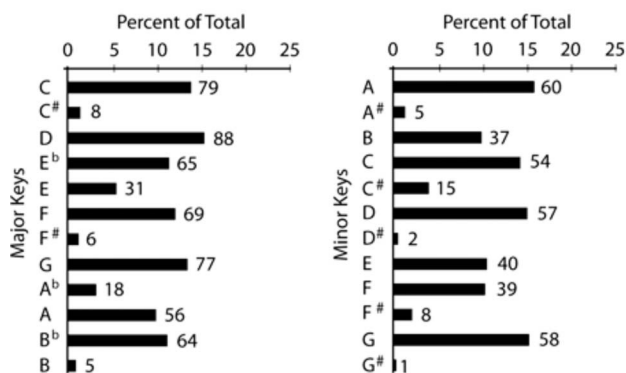


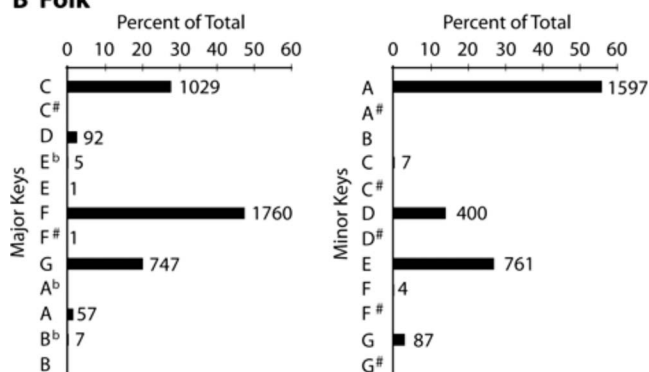
FIG. 1. The major and minor scales in Western music, shown on a piano keyboard (the abbreviations follow those in Table IA). The diatonic major scale contains only major intervals, whereas the all three diatonic minor scales substitute a minor interval at the third scale degree. Two also substitute a minor interval at the sixth scale degree, and one a minor interval at the seventh scale degree (the melodic minor scale is shown ascending; when descending it is identical to the natural minor). Thus the formal differences between the major and minor scales are in the third, sixth, and seventh scale degrees; the first, second, fourth, and fifth scale degrees are held in common.

TABLE II. The distribution of the key signatures of the major and minor melodies we compiled from Barlow and Morgenstern's Dictionary of Musical Themes (A) and the Finnish Folk Song Database (B). The number of melodies analyzed in each key is indicated to the right of each bar.

## A Classical



## B Folk



melodies in MIDI format comprising 7–69 notes. From these, we extracted 971 initial themes from works that were explicitly identified in the title as major or minor (themes from later sections were often in another key). The mean number of notes in the themes analyzed was 19.9 (SD = 7.5). To ensure that each initial theme corresponded to the key signature in the title, we verified the key by inspection of the score for accidentals; 29 further themes were excluded on this basis. Applying these criteria left 942 classical melodies for analysis of which 566 were major and 376 minor. The distribution of key signatures for these melodies is shown in Table IIA.

To ensure that our conclusions were not limited to classical music, we also analyzed major and minor melodies in the Finnish Folk Song Database (Eerola and Toivianen, 2004a). These melodies are from the traditional music of that region published in the late 19th century and first half of the 20th century (many of the themes derive from songs composed in earlier centuries). The full database contains 8614 songs in MIDI format comprising 10–934 notes each annotated by key and designated as major or minor. To make our analysis of folk songs as comparable as possible to the analysis of classical melodies and to avoid modulations in later sections of the pieces, we excluded all songs comprising more than 69 notes [the maximum number in the classical database melodies; the mean number of notes for the Finnish songs we analyzed was 39.0 (SD = 13.2)]. We also excluded

any songs that were annotated as being in more than one key or that were polyphonic. Applying these criteria left 6555 melodies for analysis of which 3699 were major and 2856 minor. The distribution of key signatures for these melodies is shown in Table IIB.

To assess the tonal differences between the major and minor melodies, the chromatic intervals represented by each melody note were determined (1) with respect to the annotated tonic of the melody and (2) with respect to the immediately following melody note. Chromatic intervals based on the tonic (referred to as “tonic intervals”) were defined as ascending with respect to the tonic by counting the number of semitones between each note in a melody and its annotated tonic (see Table IA); tonic intervals larger than an octave were collapsed into a single octave. Chromatic Intervals between melody notes (referred to as “melodic intervals”) were determined by counting the number of semitones between each note in a melody and the next. The number of occurrences of each chromatic interval was tabulated separately for tonic and melodic intervals in major and minor classical and folk music using MIDI toolbox version 1.0.1 (Eerola and Toivianen, 2004b) for MATLAB version R2007a (Mathworks Inc., 2007). The associated frequencies of occurrence were then calculated as percentages of the total number of intervals counted. For example, the frequency of tonic major thirds in classical major melodies was determined by dividing the total number of tonic major thirds in these melodies by the total number of all tonic intervals in these melodies.

The MIDI melodies were coded in equal tempered tuning; however, we converted all intervals to just intonation for analysis since this tuning system is generally considered more “natural” than equal temperament (see Sec. IV).

## C. Recording and analysis of voiced speech

Single word utterances were recorded from ten native speakers of American English (five females) ranging in age from 18–68 and without significant speech or hearing pathology. The participants gave informed consent, as required by the Duke University Health System. Monologs were also recorded from ten speakers (three of the original participants plus seven others; five females). Single words enabled more accurate frequency measurements, whereas the monologs produced more typical speech. All speech was recorded in a sound-attenuating chamber using an Audio-Technica AT4049a omni-directional capacitor microphone and a Marantz PMD670 solid-state digital recorder (Martel Electronics, Yorba Linda CA). Recordings were saved to a Sandisk flash memory card in .wav format at a sampling rate of 22.05 kHz and transferred to an Apple PowerPC G5 computer for analysis. The background noise in the recording chamber was measured in one-third-octave bands with center frequencies from 12.5 Hz to 20 kHz, using a Larson–Davis System 824 Real Time Analyzer averaging over periods of 10 s. The noise level was less than NC-15 for frequencies up to and including 500 Hz, falling to ~NC-25 by 5 kHz (NC-20 is a typical specification for an empty concert hall).

The spectra were analyzed using the “to pitch” autocor-



relation algorithm (Boersma 1993), the “to formant” LPC algorithm (Press *et al.*, 1992), and the “get power” algorithm in PRAAT version 5.0.43 (Boersma and Weenik, 2008). PRAAT default settings were used for both pitch and formant estimates, with the exception of the pitch ceiling which was set at 500 rather than 600 Hz (for pitch: floor=75 Hz, ceiling=500 Hz, time step=10 ms, window length=40 ms; for formants: number=5, ceilings for male/female =5 kHz/5.5 kHz, respectively, window length=50 ms; since there is no standard window length for power analyses we determined power every 10 ms using a 10 ms window).

*Single words.* Ten words that each had a different vowel embedded between the consonants /b/ and /d/ (i.e., bead, bid, bed, bad, bod, bud, booed, bawd, bird, and “bood,” the last pronounced like “good”) were repeated by participants. These vowels (/i, ɪ, ε, æ, ʌ, u, ɔ, ɜ, ʊ/) were chosen as a representative sample of voiced speech sounds in English; the consonant framing (/b·d/) was chosen because it maximizes vowel intelligibility (Hillenbrand and Clark, 2000). Each participant repeated each word seven times so that we could analyze the central five utterances, thus avoiding onset and offset effects. This sequence was repeated four times (twice in each emotional condition) using differently ordered lists of the words, pausing for 30 s between the recitations of each word list. Participants were instructed to utter the words as if they were excited and happy, or conversely as if they were subdued and sad. The quality of each subject’s performance was monitored remotely to ensure that the speech was easily recognized as either excited or subdued. The fundamental frequency values were checked for accuracy by comparing the pitch track in PRAAT with the spectrogram. In total, 1000 examples of single word vowel sounds in each experimental condition were included in the database (i.e., 100 from each speaker).

A PRAAT script was used to automatically mark the pauses between each word; vowel identifier and positional information were then inserted manually for each utterance (the script can be accessed at [http://www.helsinki.fi/~lennes/praat-scripts/public/mark\\_pauses.praat](http://www.helsinki.fi/~lennes/praat-scripts/public/mark_pauses.praat)). A second script was written to extract the fundamental frequency (using “to pitch”) and the frequencies of the first and second formants (using “to formant”) from a 50 ms segment at the mid-point of each vowel utterance.

*Monologs.* The participants read five monologs with exciting content and five monologs with subdued content (Table III). Each monolog was presented for reading in a standard manner on a computer monitor and comprised about 100 syllables with an approximately equal distribution of the ten different American English vowels examined in the single word analysis. The instructions given the participants were to utter the monologs with an emotional coloring appropriate to the content after an initial silent reading. Performance quality was again monitored remotely.

Analysis of the monologs was complicated by the fact that natural speech is largely continuous; thus the “mark pauses” script could not be used. Instead we wrote another PRAAT script to extract the fundamental frequency, first and second formant frequencies, and amplitude values using a time step of 10 ms and the PRAAT functions listed above

TABLE III. Examples of the excited and subdued monologs read by the speakers.

(A) Excited
I won the lottery and I still can’t believe it! I always bought a ticket but you never expect to win. I don’t know what I’m going to do with ten million dollars but I’m sure going to have fun finding out. I’ve had some good luck in my day, but this tops it all.
(B) Subdued
The papers finalizing the divorce came today. Sometimes I think that it is better this way but still I am torn between hurting the kids and staying in a hopeless marriage, there is really no solution. They are still so young and I know that this will hurt them a lot.

(amplitude was obtained by taking the square root of the values returned by “get power”). In order to avoid the inclusion of silent intervals, time points where the amplitude was less than 5% of a speaker’s maximum were rejected. To avoid including unvoiced speech segments, time points at which no fundamental frequency was identified were also removed (as the to pitch function steps through the speech signal it assigns time points as either voiced or unvoiced depending on the strength of identified fundamental frequency candidates; for details, see Boersma, 1993). On average, 32% of the monolog recordings was silence and 13% was unvoiced; thus 55% of the material was voiced and retained. Each monolog was spoken over 10–20 s and yielded approximately 380 voiced data points for analysis.

## D. Comparison of speech and musical spectra

Spectral comparisons were based on fundamental frequency and frequency ratios; these two acoustic features were chosen because of the critical roles they play in the perception of both voiced speech sounds and musical intervals. In speech, fundamental frequency carries information about the sex, age, and emotional state of a speaker (Hollien, 1960; Crystal, 1990; Protopapas and Lieberman, 1996; Banse and Scherer, 1996; Harrington *et al.*, 2007); frequency ratios between the first and second formants (F1, F2) differentiate particular vowel sounds, allowing them to be understood across speakers with anatomically different vocal tracts (Delattre, 1952; Pickett *et al.*, 1957; Petersen and Barney, 1962; Crystal, 1990; Hillenbrand *et al.*, 1995). In music, the fundamental frequencies of the notes carry the melody; the frequency ratios between notes in the melody and the tonic define the intervals and provide the context that determines whether the composition is in a major or minor mode.

*Comparison of fundamental frequencies.* The fundamental frequency of each voiced speech segment was determined as described [Fig. 2(a)]. The comparable fundamental frequency of a musical interval, however, depends on the relationship between two notes. The harmonics of two notes can be thought of as the elements of a single harmonic series with a fundamental defined by their greatest common divisor [Fig. 2(b)]. Accordingly, for each of the intervals that distinguished major and minor music in our databases (see Fig. 1 and below), we calculated the frequency of the greatest common divisor of the relevant notes. For tonic intervals these notes were the melody note and its annotated tonic, and for

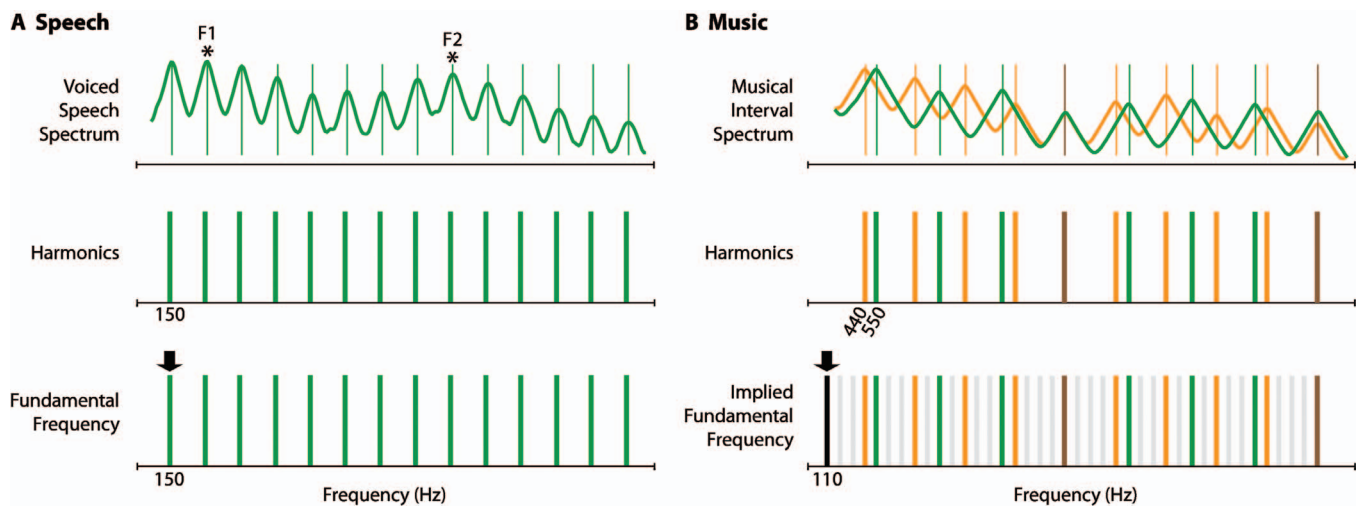


FIG. 2. The harmonic structure of speech sounds and musical intervals. (A) The spectrum of a voiced speech sound comprises a single harmonic series generated by the vibration of the vocal folds (the vertical green lines indicate the loci of harmonic peaks); the relative amplitude of the harmonics is modulated by the resonance of the rest of the vocal tract, thus defining speech formants (asterisks indicate the harmonic peaks of the first two formants, F1 and F2). The voiced speech segment shown here as an example was taken from the single word database and has a fundamental frequency of 150 Hz (black arrow in the lower panel). (B) The spectra of musical intervals entail two harmonic series, one from each of the two relevant notes (see text). The example in the upper panel shows the superimposed spectra of two musical notes related by a major third (the harmonic series of the lower note is shown in orange, and the harmonic series of the higher note in green, and the harmonics common to both series in brown). Each trace was generated by averaging the spectra of 100 recordings of tones played on an acoustic guitar with fundamentals of  $\sim 440$  and  $\sim 550$  Hz, respectively; the recordings were made under the same conditions as speech (see Sec. II). The implied fundamental frequency (black arrow in the lower panel) is the greatest common divisor of the two harmonic series.

melodic intervals these notes were the two adjacent notes in the melody. Differences between the distributions of fundamental frequencies in excited and subdued speech, and between the distributions of implied fundamental frequencies in major and minor music, were evaluated for statistical significance by independent two-sample *t*-tests.

The perceptual relevance of implied fundamental frequencies at the greatest common divisor of two tones is well documented (virtual pitch; reviewed in Terhardt, 1974). However, in order for an implied fundamental to be heard the two tones must be presented simultaneously or in close succession (Hall and Peters, 1981; Grose *et al.*, 2002). These criteria are not met by the tonic intervals we examined here because for many melody notes the tonic will not have been sounded for some time in the melody line. Instead the justification for using implied fundamentals with tonic intervals depends on the tonic's role in providing the tonal context for appreciating the other melody notes (Randel, 1986; Aldwell and Schacter, 2003). Each note in a melody is perceived in the context of its tonic regardless of whether or not the tonic is physically simultaneous. This is evident from at least two facts: (1) if the notes in a melody were not perceived in relation to the tonic, there would be no basis for determining the interval relationships that distinguish one mode from another, making it impossible to distinguish major compositions from minor ones (Krumhansl, 1990; Aldwell and Schacter, 2003); (2) because a note played in isolation has no affective impact whereas a note played in the context of a melody does, the context is what gives individual notes their emotional meaning (Huron, 2006). Thus, the fact that we can hear the difference between major and minor melodies and that we are affected by each in a characteristic way indicates that each note in a melody is, in a very real sense, heard in the context of a tonic.

*Comparison of speech formants and musical ratios.* The frequency ratios of the first two formants in excited and subdued speech were compared with the frequency ratios of the intervals that specifically distinguish major and minor music. To index F1 and F2 we used the harmonic nearest the peak formant value given by the linear predictive coding (LPC) algorithm utilized by PRAAT's "to formant" function, i.e., the harmonic with the greatest local amplitude (in a normal voiced speech sound there may or may not be a harmonic power maximum at the LPC peak itself; Fant 1960; Press *et al.*, 1992). The frequency values of the harmonics closest to the LPC peaks were determined by multiplying the fundamental frequency by sequentially increasing integers until the difference between the result and the LPC peak was minimized. The ratios of the first two formants were then calculated as  $F2/F1$  and counted as chromatic if they were within 1% of the just intonation ratios in Table IA. Differences in the prevalence of specific formant ratios in excited and subdued speech were evaluated for statistical significance using chi-squared tests for independence. The analysis focused on F1 and F2 because they are the most powerful resonances of the vocal tract and because they are necessary and sufficient for the discrimination of vowel sounds (Delattre, 1952; Pickett *et al.*, 1957; Rosner and Pickering, 1994). Other formants (e.g., F3 and F4) are also important in speech perception, but are typically lower in amplitude and not critical for the discrimination of vowels (op cit.).

### III. RESULTS

#### A. Intervals in major and minor music

The occurrence of different chromatic intervals among the tonic and melodic intervals in major and minor classical and folk music is shown in Table IV. As expected from the

TABLE IV. Frequency of occurrence of chromatic intervals in major and minor Western classical and Finnish folk music. (A) Tonic intervals; defined as the number of semitones between a melody note and its tonic. (B) Melodic intervals; defined as the number of semitones between adjacent melody notes. The preponderance of small intervals in (B) is in agreement with previous studies (Vos and Troost, 1989). The intervals that distinguish major and minor music are underlined (dashed-lines indicate intervals with less marked contributions).

Major melodies			Minor melodies		
Intervals	Classical (%)	Folk (%)	Intervals	Classical (%)	Folk (%)
(A) Tonic intervals					
Unison/octave	19.9	20.4	Unison/octave	19.4	19.1
Minor second	0.4	0.1	Minor second	0.6	0.2
Major second	12.8	15.9	Major second	13.0	19.6
Minor third	0.8	0.0	Minor third	15.8	15.6
Major third	18.2	16.8	Major third	0.7	0.2
Perfect fourth	10.6	9.5	Perfect fourth	10.5	10.1
Tritone	1.1	0.4	Tritone	1.7	0.2
Perfect fifth	19.1	19.6	Perfect fifth	20.3	19.9
Minor sixth	0.4	0.0	Minor sixth	7.9	1.6
Major sixth	8.4	8.9	Major sixth	1.3	2.9
Minor seventh	0.6	0.2	Minor seventh	3.4	7.5
Major seventh	7.7	8.1	Major seventh	5.4	3.1
(B) Melodic intervals					
Unison	10.7	24.9	Unison	11.4	24.0
Minor second	20.8	13.3	Minor second	28.2	19.1
Major second	36.1	29.9	Major second	27.6	27.6
Minor third	9.4	11.6	Minor third	10.4	12.1
Major third	6.9	8.6	Major third	5.4	6.2
Perfect fourth	7.6	7.4	Perfect fourth	7.2	7.1
Tritone	0.4	0.2	Tritone	1.1	0.0
Perfect fifth	2.7	2.0	Perfect fifth	3.2	2.5
Minor sixth	1.1	0.6	Minor sixth	1.8	0.8
Major sixth	1.3	0.9	Major sixth	1.2	0.1
Minor seventh	0.4	0.3	Minor seventh	0.4	0.1
Major seventh	0.1	0.0	Major seventh	0.2	0.0
Octave	1.3	0.3	Octave	1.2	0.2
Larger	1.0	0.0	Larger	0.9	0.0

formal structure of major and minor scales (see Fig. 1) as well as musical practice, the most salient empirical distinction between major and minor music is the tonic interval of the third scale degree. In both the classical and folk databases major thirds made up 16%–18% of the intervals in major melodies and less than 1% of the intervals in minor melodies; this pattern was reversed for minor thirds, which comprised less than 1% of the intervals in major melodies and about 15% of the intervals in minor melodies. Tonic intervals of the sixth and seventh scale degrees also distinguish major and minor music, but less obviously. These intervals are only about half as prevalent in music as thirds, and their distribution in major versus minor music is less differentiated. There were no marked differences between major and minor music among the tonic intervals held in common by major and minor scales (unison/octave, perfect fifth, major second, perfect fourth; see Fig. 1 and Table IVA).

Notable differences between major and minor music were also found for melodic seconds. In the classical database, there were ~8% more melodic major seconds in major melodies than minor melodies, and ~7% more melodic minor seconds in minor melodies than major melodies. In the folk database this same pattern was apparent, with major melodies containing ~2% more melodic major seconds than minor melodies, and minor melodies containing ~6% more melodic minor seconds than major melodies (Table IVB). The prevalence of the other melodic intervals was similar across major and minor music, differing by 1% or less for all except major thirds in folk music, which were 2.4% more prevalent in major melodies than minor melodies.

The average duration of notes in major and minor melodies was comparable, being within 30 ms in the classical database and within 5 ms in the folk database. The overall difference in the mean pitch height of major and minor melo-



dies in the classical and folk databases was also small, with minor melodies being played an average of 10 Hz higher than major melodies.

This empirical analysis of a large number of melodies in both classical and folk genres thus documents the consensus that the primary tonal distinction between major and minor music with respect to tonic intervals is the prevalence of major versus minor thirds, with smaller differences in the prevalence of major and minor sixths and sevenths (see Sec. IV). With respect to melodic intervals, the only salient distinction between major and minor melodies is the prevalence of major versus minor seconds, major music being characterized by an increased prevalence of major seconds and minor music by an increased prevalence of minor seconds.

## B. Comparison of fundamental frequencies

Figure 3 shows the distributions of fundamental frequencies for individual speakers uttering speech in an excited compared to a subdued manner. In agreement with previous observations (Banse and Scherer 1996; Juslin and Laukka, 2003; Scherer 2003; Hammerschmidt and Jurgens, 2007), the fundamental frequency of excited speech is higher than that of subdued speech. On average, the mean fundamental frequency of excited speech exceeded that of subdued speech by 150 Hz for females and 139 Hz for males in the single word condition, and by 58 Hz for females and 39 Hz for males in the monologue condition.

Figure 4 shows the distributions of the implied fundamental frequencies for tonic thirds, sixths, and sevenths and melodic seconds in major and minor melodies. As expected from the ratios that define major and minor thirds (5:4 and 6:5, respectively; see Sec. IV) and their different prevalence in major and minor music (see Table IVA), the implied fundamentals of tonic thirds are significantly higher in major music than in minor music for both genres examined. In classical music, the mean implied fundamental of thirds in major melodies is higher than that in minor melodies by 21 Hz; in folk music, the mean implied fundamental of thirds in major melodies is higher than that in minor melodies by 15 Hz. The pattern for tonic sixths is similar, with the mean implied fundamental of sixths in major melodies being higher than that in minor melodies by 46 Hz in classical music, and by 13 Hz in folk music, despite the occurrence of more major than minor sixths in minor folk music. The pattern for sevenths in major and minor music, however, is different than that of thirds and sixths. Whereas in folk music the mean implied fundamental of sevenths in major melodies is a little higher than that in minor music (by  $\sim 2$  Hz), in classical music the mean implied fundamental of sevenths in major melodies is actually lower than that in minor melodies (by  $\sim 5$  Hz; see Sec. IV). The differences between the implied fundamental frequencies of melodic seconds in major versus minor melodies follow the same pattern as tonic thirds and sixths with a higher mean implied fundamental in major music than in minor music in both musical databases. However, as with tonic sevenths the differences between the mean implied fundamentals of seconds in major and minor melodies were relatively small ( $\sim 3$  and  $\sim 2$  Hz for classical and

folk music, respectively; all of these differences are highly significant with  $p$ -values  $< 0.0001$  or less).

An additional consideration in the comparison of fundamental frequencies in speech and the implied fundamental frequencies of musical intervals is whether they are within the same range. In the speech database,  $\sim 97\%$  of fundamental frequencies were between 75 and 300 Hz. In the music database, the percentage of implied fundamentals within this range was different for each of the intervals considered. For tonic thirds and sixths,  $\sim 70\%$  and  $\sim 56\%$  of implied fundamentals were in this range, respectively. However, only  $\sim 1\%$  of the implied fundamentals of tonic sevenths and  $\sim 7\%$  of those of melodic seconds were within this range (see Fig. 3). Thus the implied fundamental frequencies of the spectra of tonic thirds and sixths, but not tonic sevenths and melodic seconds, conform to the frequency range of the fundamental frequencies in speech.

These results show that the implied fundamentals of tonic thirds and sixths but not other intervals that empirically distinguish major and minor music parallel to the differences in the fundamental frequencies of excited and subdued speech.

## C. Comparison of formant and musical ratios

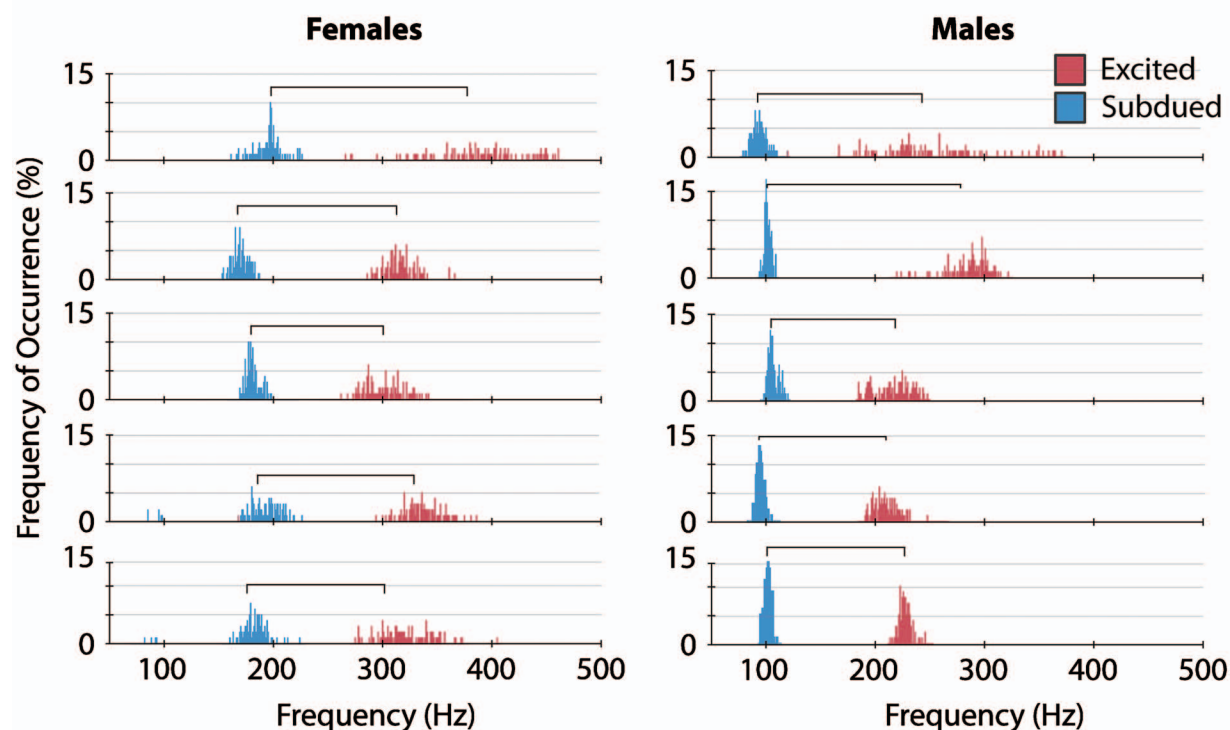
The distributions of F2/F1 ratios in excited and subdued speech spectra are shown in Fig. 5. In both single words and monologs, excited speech is characterized by a greater number of major interval ratios and relatively few minor interval ratios, whereas subdued speech is characterized by relatively fewer major interval ratios and more minor interval ratios. Thus in the single word database, F2/F1 ratios corresponding to major seconds, thirds, sixths, and sevenths made up  $\sim 36\%$  of the ratios in excited speech, whereas ratios corresponding to minor seconds, thirds, sixths, and sevenths were entirely absent. In subdued speech only  $\sim 20\%$  of the formant ratios corresponded to major seconds, thirds, sixths, and sevenths, whereas  $\sim 10\%$  of the ratios corresponded to minor seconds, thirds, sixths, and sevenths. The same trend is evident in the monolog data; however, because the overlap of the distributions of excited and subdued fundamentals is greater, the differences are less pronounced (see Fig. 3 and Sec. IV).

These parallel differences between the occurrence of formant ratios in excited and subdued speech and the ratios of the musical intervals that distinguish major and minor melodies provide a further basis for associating the spectra of speech in different emotional states with the spectra of intervals that distinguish major and minor music.

## IV. DISCUSSION

Although the tonal relationships in music are only one determinant of its affective impact—the emotional influence of tonality in melodies competes with the effects of intensity, tempo, and rhythm among other factors—they are clearly consequential, as indicated by the distinct affective qualities of major and minor music. Despite the use of major and minor tonalities for affective purposes in Western music for at least the last 400 years (Zarlino, 1558), the reason for their

## A Single Words



## B Monologues

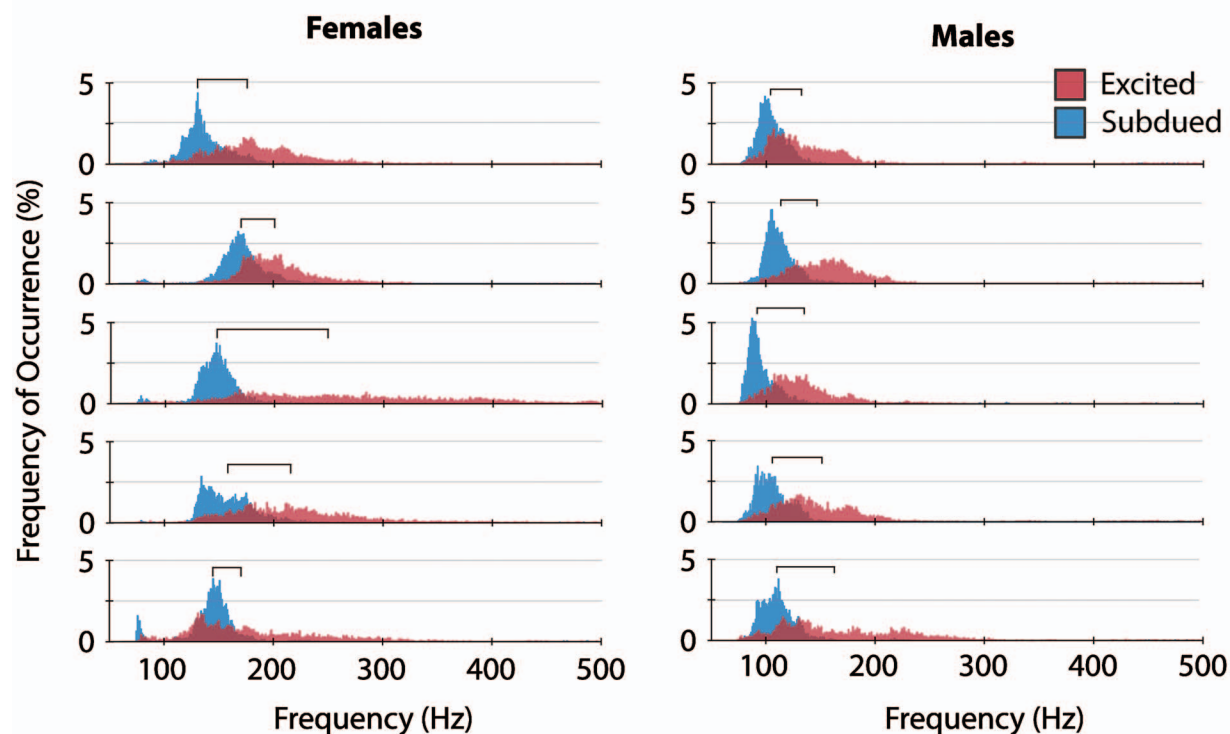


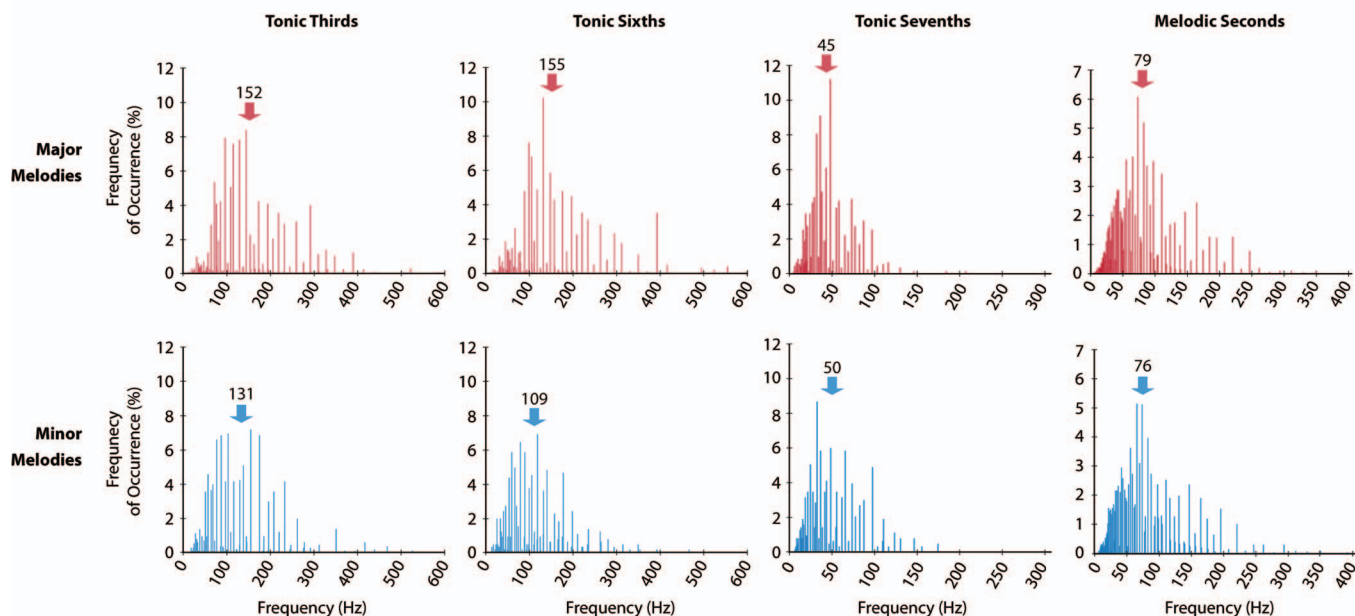
FIG. 3. The fundamental frequencies of excited (red) and subdued (blue) speech segments for individual male and female speakers derived from (A) single word utterances and (B) monolog recordings; brackets indicate means. The differences between the two distributions are significant for each speaker ( $p < 0.0001$  or less in independent two-sample t-tests). The difference between the mean fundamentals of the excited and subdued distributions is also significant across speakers ( $p < 0.0001$  for single words, and  $< 0.01$  for monologs in dependent t-tests for paired samples).

different emotional quality is not known. We thus asked whether the affective character of major versus minor music might be a consequence of associations made between spec-

tral qualities of the intervals that specifically distinguish major and minor compositions and spectral qualities of speech uttered in different emotional states.



## A Western Classical Music



## B Finnish Folk Music

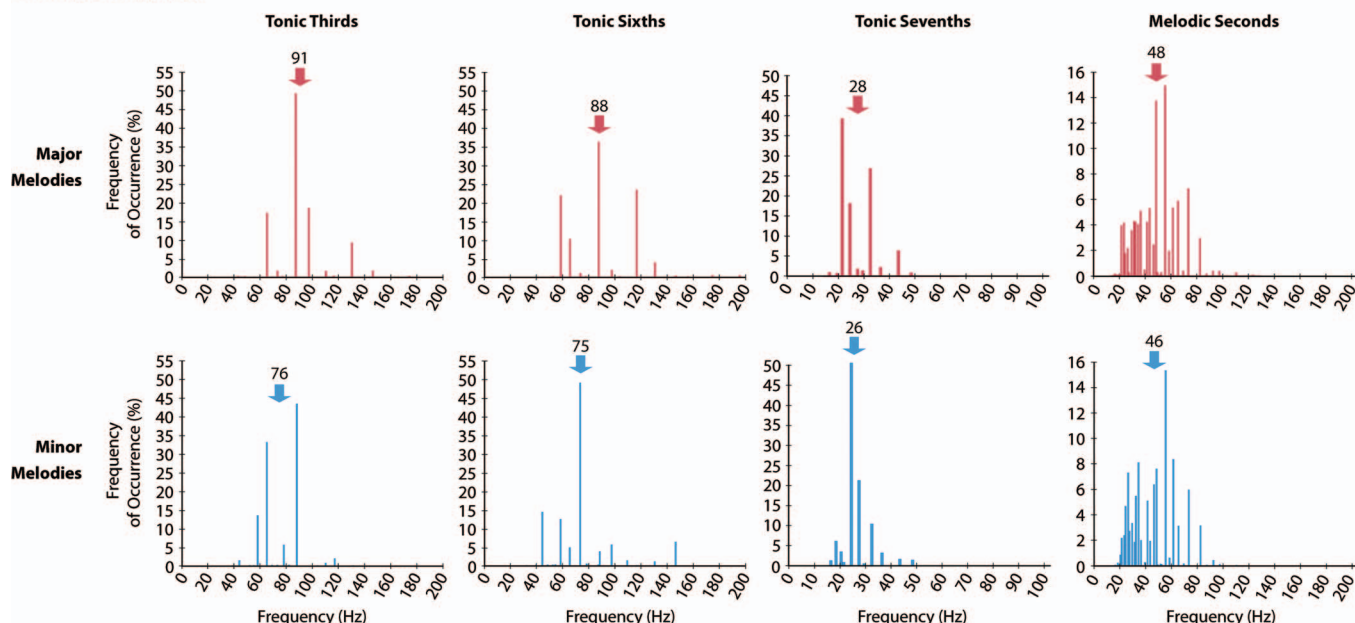


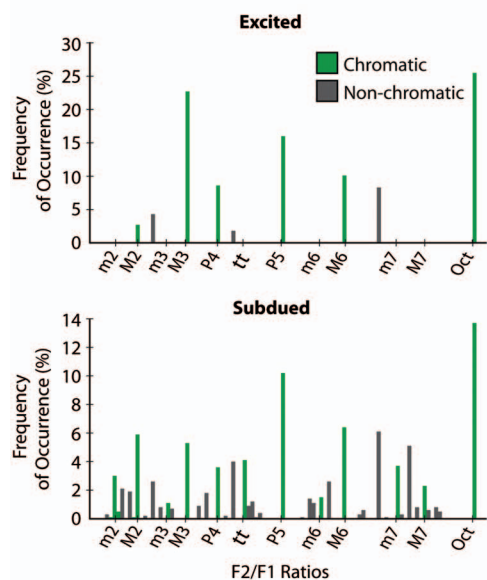
FIG. 4. The implied fundamental frequencies of tonic thirds, sixths, and sevenths and melodic seconds in major and minor melodies from Western classical music (A) and Finnish folk music (B). Arrows indicate the mean implied fundamental frequency values for each distribution. The sparseness of the tonic interval data in (B) is a consequence of fewer different key signatures in our folk music sample compared to our sample of classical music (see Table II), as well as the fact that the classical melodies span more octaves than the folk melodies ( $\sim 6$  octaves vs  $\sim 3$  octaves). Differences between the distributions of implied fundamental frequencies for major and minor melodies are statistically significant for each of the intervals compared ( $p < 0.0075$  or less in independent two-sample t-tests).

## A. Empirical differences between major and minor music

Analysis of a large number of classical and folk melodies shows that the principal empirical distinction between major and minor music is the frequency of occurrence of tonic major and minor thirds: nearly all tonic thirds in major melodies are major, whereas this pattern is reversed in minor melodies (see Table IVA). The different distributions of tonic major and minor sixths and sevenths also contribute to the tonal distinction of major and minor music, but less promi-

nently. Although major melodies almost never contain tonic minor sixths and sevenths, melodies in a minor key often include tonic major sixths and sevenths, presumably because the harmonic and melodic versions of the minor scales are often used (see Fig. 1). An additional distinction between major and minor music is the prevalence of melodic major and minor seconds. Compared to minor melodies, major melodies contain more melodic major seconds, and, compared to major melodies, minor melodies contain more melodic minor seconds (see Table IVB).

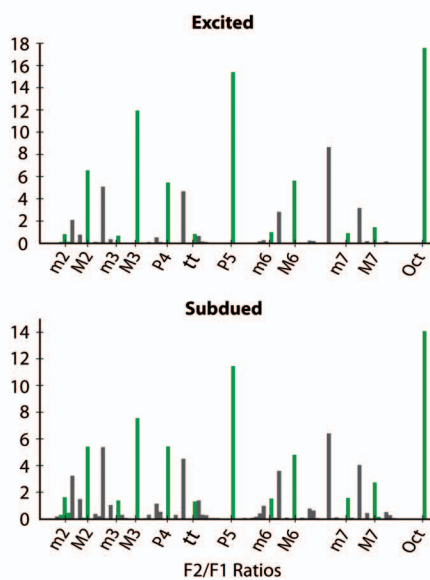
## A Single Words



## B Single Words

Intervals	Excited	Subdued	p-value
Unison/Octave	25.5 %	13.7 %	< .0001
Minor Second	0.0 %	3.5 %	< .0001
Major Second	2.7 %	5.8 %	< .001
Minor Third	0.0 %	1.1 %	< .001
Major Third	22.7 %	5.3 %	< .0001
Perfect Fourth	8.6 %	3.6 %	< .0001
Tritone	0.0 %	4.1 %	< .0001
Perfect Fifth	16.0 %	10.2 %	< .0005
Minor Sixth	0.0 %	1.5 %	< .0001
Major Sixth	10.1 %	6.4 %	< .01
Minor Seventh	0.0 %	3.7 %	< .0001
Major Seventh	0.0 %	2.3 %	< .0001
Non-Chromatic	14.4 %	38.7 %	< .0001

## C Monologues



## D Monologues

Intervals	Excited	Subdued	p-value
Unison/Octave	17.6 %	14.1 %	< .0001
Minor Second	1.0 %	2.1 %	< .0001
Major Second	6.6 %	5.4 %	< .0001
Minor Third	0.7 %	1.4 %	< .0001
Major Third	12.0 %	7.6 %	< .0001
Perfect Fourth	5.5 %	5.4 %	= .82
Tritone	0.8 %	1.3 %	< .0001
Perfect Fifth	15.4 %	11.5 %	< .0001
Minor Sixth	1.0 %	1.5 %	< .0001
Major Sixth	5.6 %	4.8 %	< .0001
Minor Seventh	0.9 %	1.6 %	< .0001
Major Seventh	1.4 %	2.7 %	< .0001
Non-Chromatic	31.6 %	40.7 %	< .0001

FIG. 5. Comparison of the ratios of the first two formants in excited and subdued speech derived from analyses of the single word and monologue databases. Ratios have been collapsed into a single octave such that they range from 1 to 2. (A) The distribution of formant ratios in excited and subdued speech from the single word database; green bars indicate ratios within 1% of chromatic interval ratios (see Table 1A); gray bars indicate ratios that did not meet this criterion. (B) The percentage of formant ratios corresponding to each chromatic interval in (A) for excited and subdued speech. (C) The same as (A), but for the monologue data. (D) The same as (B), but for the monologue data in (C). *p*-values for each interval were calculated using the chi-squared test for independence, with expected values equal to the mean number of occurrences of an interval ratio across excited and subdued speech. Intervals empirically determined to distinguish major and minor music are underlined (dashed-lines indicate intervals with less marked contributions).

## B. Comparison of speech and music spectra

Speech uttered in an excited compared to a subdued manner differs in pitch, intensity, tempo, and a number of other ways (Bense and Scherer 1996; Scherer 2003; Hammerschmidt and Jurgens, 2007). With respect to tonality, however, the comparison of interest lies in the spectral features of voiced speech and musical intervals (see Introduction).

The results we describe show that differences in both fundamental frequency and formant relationships in excited versus subdued speech spectra parallel spectral features that distinguish major and minor music. When a speaker is excited, the generally increased tension of the vocal folds raises the fundamental frequency of vocalization; conversely when speakers are in a subdued state decreased muscular tension lowers the fundamental frequency of vocalization (see Fig. 3). In music two factors determine the frequency of a musical interval's implied fundamental: (1) the ratio that defines the interval and (2) the pitch height at which the interval is played. The defining ratios of minor seconds, thirds, and sixths (16:15, 6:5 and 8:5, respectively) yield smaller greatest common divisors than the defining ratios of major seconds, thirds, and sixths (9:8, 5:4, and 5:3); thus minor seconds, thirds, and sixths played at the same pitch height as

their major counterparts will always have lower implied fundamentals. In our musical databases, the average pitch height of minor melodies was somewhat higher than major melodies, thus reducing the difference between the implied fundamentals of these major and minor intervals (see Sec. III). Despite these effects, the mean implied fundamentals of tonic thirds and sixths and melodic seconds in major music are still higher than those in minor music. However, for melodic seconds, the relatively small size of this difference ( $\sim 2\text{--}3$  Hz) and the fact that most of their implied fundamentals ( $\sim 93\%$ ) fall below the range of fundamental frequencies in speech make associations with the spectra of excited and subdued speech on these grounds less likely. In contrast to the defining ratios of seconds, thirds, and sixths, the defining ratio of the minor seventh (9:5) yields a *larger* greatest common divisor than the defining ratio of the major seventh (16:9), making its implied fundamental at a given pitch height higher than its major counterpart. This reasoning indicates why the pattern observed for tonic thirds and sixths and melodic seconds is not also apparent for tonic sevenths. Furthermore, as with melodic seconds, nearly all of the implied fundamentals of tonic sevenths ( $\sim 99\%$ ) fall below the range of the fundamentals in speech.

The difference between the fundamental frequencies of

excited and subdued speech also affects the prevalence of specific formant ratios. Given the same voiced speech sound, the positions of the first and second formants are relatively stable between excited and subdued speech (as they must be to allow vowel phonemes to be heard correctly). The higher fundamental frequencies of excited speech, however, increase the frequency distance between harmonics, causing lower harmonics to underlie the first and second formants. As a result the F2/F1 ratios in excited speech tend to comprise smaller numbers and thus more often represent musical intervals defined by smaller number ratios. Conversely, the lower fundamental frequencies in subdued speech decrease the distance between harmonics, causing higher harmonics to underlie the formants. Thus the F2/F1 ratios in subdued speech tend to comprise larger numbers, which more often represent musical intervals defined by larger number ratios. Intervals whose defining ratios contain only the numbers one through five (octaves, perfect fifths, perfect fourths, major thirds, and major sixths) were typically more prevalent in the F2/F1 ratios of excited speech, whereas intervals with defining ratios containing larger numbers (all other chromatic intervals) were more prevalent in the F2/F1 ratios of subdued speech (see Table I and Fig. 5). The only exceptions to this rule were the prevalences of major seconds and perfect fourths in the F2/F1 ratios of the monolog recordings, which were slightly higher in excited speech than in subdued speech or not significantly different, respectively. Presumably, these exceptions result from the greater overlap of fundamental frequencies between excited and subdued speech in the monologues (see Fig. 3).

Thus, based on these comparisons, differences in the spectra of excited and subdued speech parallel differences in the spectra of major and minor tonic thirds and sixths, but not of major and minor melodic seconds and tonic sevenths.

### C. Melodic seconds and tonic sevenths

Although the empirical prevalence of melodic seconds and tonic sevenths also distinguishes major and minor music (see Table IV), as described in Sec. IV B the spectral characteristics of these intervals do not parallel those in excited and subdued speech. The mean implied fundamental frequency of tonic sevenths in major music is lower than in minor music, and, although the mean implied fundamental frequency of melodic seconds in major music is higher than in minor music the difference is small. Moreover, the majority of implied fundamentals for these intervals are below the range of fundamental frequencies in speech. With respect to formant ratios, there are fewer minor seconds and sevenths in excited speech than in subdued speech, but this was also true of major seconds and sevenths.

These results accord with music theory. Unlike tonic thirds and sixths, melodic seconds and tonic sevenths are not taken to play a significant role in distinguishing major and minor music (Aldwell and Schachter, 2003). Rather these intervals are generally described as serving other purposes, such as varying the intensity of melodic motion in the case of seconds and creating a sense of tension that calls for resolution to the tonic in the case of sevenths (op cit.). This

interpretation is supported by the fact that major and minor melodic seconds and tonic sevenths are commonplace in both major and minor music (see Table IV).

### D. Just intonation vs equal temperament

Most of the music analyzed here will have been performed and heard in equal tempered tuning, as is almost all popular music today; given that our hypothesis is based on associations between speech and music, our decision to use just intoned rather than equally tempered ratios to define the musical intervals requires justification.

Just intonation is based on ratios of small integers and thus is readily apparent in the early harmonics of any harmonic series, voiced speech included (Rossing, 1990). To relate the spectral structure of voiced speech to the spectra of musical intervals, it follows that just intoned ratios are the appropriate comparison. Equal tempered tuning is a compromise that allows musicians to modulate between different key signatures without retuning their instruments while retaining as many of the perceptual qualities of just intoned intervals as possible. The fact that the differences introduced by equal tempered tuning are acceptable to most listeners implies that either system is capable of associating the harmonic characteristics of music and speech.

Furthermore, the spectral comparisons we made do not strictly depend on the use of just intoned ratios. With respect to implied fundamental frequencies, the phenomenon of virtual pitch is robust and does not depend on precise frequency relations (virtual pitches can be demonstrated with equally tempered instruments such as the piano; Terhardt *et al.*, 1982). With respect to formant ratios, if equal tempered tunings had been used to define the intervals instead of just intonation tunings, the same results would have been obtained by increasing the window for a match from 1% to 2%. Finally, retuning the major and minor thirds and sixths for the 7497 melodies we examined, resulted in a mean absolute frequency difference of only 3.5 Hz.

### E. Why thirds are preeminent in distinguishing major and minor music

A final question is why the musical and emotional distinction between major and minor melodies depends primarily on tonic thirds, and how this fact aligns with the hypothesis that associations made between the spectral characteristics of music and speech are the basis for the affective impact of major versus minor music. Although our analysis demonstrates similarities between the spectra of major and minor thirds and the spectra of excited and subdued speech, it does not indicate why thirds, in particular, provide the critical affective difference arising from the tonality of major and minor music. One plausible suggestion is that among the intervals that differentiate major and minor tone collections, thirds entail the lowest and thus the most powerful harmonics. Accordingly, of the distinguishing intervals, thirds are likely to be the most salient in the spectra of both voiced speech sounds and musical tones.



## V. CONCLUSION

In most aspects of music—e.g., intensity, tempo, and rhythm—the emotional quality of a melody is conveyed at least in part by physical imitation in music of the characteristics of the way a given emotion is expressed in human behavior. Here we asked whether the same principle might apply to the affective impact of the different tone collections used in musical compositions. The comparisons of speech and music we report show that the spectral characteristics of excited speech more closely reflect the spectral characteristics of intervals in major music, whereas the spectral characteristics of subdued speech more closely reflect the spectral characteristics of intervals that distinguish minor music. Routine associations made between the spectra of speech uttered in different emotional states and the spectra of thirds and sixths in major and minor music thus provide a plausible basis for the different emotional effects of these different tone collections in music.

## ACKNOWLEDGMENTS

We are grateful to Sheena Baraton, Nigel Barella, Brent Chancellor, Henry Greenside, Yizhang He, Dewey Lawson, Kevin LaBar, Rich Mooney, Deborah Ross, David Schwartz, and Jim Voyvodic for useful comments and suggestions. We especially wish to thank Dewey Lawson for his help measuring the acoustical properties of the recording chamber, to Deborah Ross help selecting and recording the monologs, to Shin Chang, Lap-Ching Keung, and Stephan Lotfi who helped organize and validate the musical databases, and to Tuomas Eerola for facilitating our access to the Finnish Folk Song Database.

Aldwell, E., and Schachter, C. (2003). *Harmony & Voice Leading*, 3rd ed. (Wadsworth Group/Thomson Learning, Belmont, CA).

Banse, R., and Scherer, K. R. (1996). "Acoustic profiles in vocal emotion expression," *J. Pers. Soc. Psychol.* **70**, 614–636.

Barlow, H., and Morgenstern, S. (1974). *A Dictionary of Musical Themes* (Crown, New York).

Bernstein, L. (1976). *The Unanswered Question: Six Talks at Harvard* (Harvard University Press, Cambridge, MA).

Boersma, P. (1993). "Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound," *Proceedings of the Institute of Phonetic Sciences*, **17**, pp. 97–110, University of Amsterdam.

Boersma, P., and Weenik, D. (2008). PRAAT: Doing phonetics by computer (Version 5.0.43), available from <http://www.fon.hum.uva.nl/praat/> (Last viewed November, 2008).

Burkholder, J. P., Grout, D., and Palisca, C. (2005). *A History of Western Music*, 7th ed. (Norton, New York).

Burns, E. M. (1999). "Intervals, scales and tuning," in *The Psychology of Music*, 2nd ed., edited by D. Deutsch (Academic, New York).

Carterette, E. C., and Kendall, R. A. (1999). "Comparative music perception and cognition," in *The Psychology of Music*, 2nd ed., edited by D. Deutsch (Academic, New York).

Cohen, D. (1971). "Palestrina counterpoint: A musical expression of unexcited speech," *Journal of Music Theory* **15**, 85–111.

Cooke, D. (1959). *The Language of Music* (Oxford University Press, Oxford, UK).

Crowder, R. G. (1984). "Perception of the major/minor distinction: Hedonic, musical, and affective discriminations," *Bull. Psychon. Soc.* **23**, 314–316.

Crystal, D. (1997). *The Cambridge Encyclopedia of Language*, 2nd ed. (Cambridge University Press, New York).

Delattre, P., Liberman, A. M., Cooper, F. S., and Gerstman, L. J. (1952).

"An experimental study of the acoustic determinants of vowel color: Observation of one- and two-formant vowels synthesized from spectrographic patterns," *Word* **8**, 195–210.

Eerola, T., and Tovianen, P. (2004a). Suomen Kasan eSävelmät (Finnish Folk Song Database), available from <http://www.jyu.fi/musica/sks/> (Last viewed November, 2008).

Eerola, T., and Toiviainen, P. (2004b). MIDI Toolbox: MATLAB Tools for Music Research (Version 1.0.1), available from <http://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/miditoolbox/> (Last viewed November, 2008).

Fant, G. (1960). *Acoustic Theory of Speech Production* (Mouton, The Hague).

Gabrielsson, A., and Lindström, E. (2001). "The influence of musical structure on emotional expression," in *Music and Emotion: Theory and Research*, 1st ed., edited by P. N. Juslin and J. A. Sloboda (Oxford University Press, New York).

Gregory, A. H., and Varney, N. (1996). "Cross-cultural Comparisons in the affective response to music," *Psychol. Music* **24**, 47–52.

Grose, J. H., Hall, J. W., III, and Buss, E. (2002). "Virtual pitch integration for asynchronous harmonics," *J. Acoust. Soc. Am.* **112**, 2956–2961.

Hall, J. W., and Peters, R. W. (1981). "Pitch for non-simultaneous harmonics in quiet and noise," *J. Acoust. Soc. Am.* **69**, 509–513.

Hammerschmidt, K., and Jurgens, U. (2007). "Acoustical correlates of affective prosody," *J. Voice* **21**, 531–540.

Harrington, J., Palethorpe, S., and Watson, C. I. (2007). "Age-related changes in fundamental frequency and formants: a longitudinal study of four speakers," in *Proceedings of Interspeech*, 2007, Antwerp.

Heinlein, C. P. (1928). "The affective characters of the major and minor modes in music," *J. Comp. Psychol.* **8**, 101–142.

Helmholtz, H. (1885). *Lehre von den tonempfindungen (On the Sensations of Tone)*, 4th ed., translated by A. J. Ellis (Dover, New York).

Hevner, K. (1935). "The affective character of the major and minor modes in music," *Am. J. Psychol.* **47**, 103–118.

Hillenbrand, J. M., and Clark, M. J. (2001). "Effects of consonant environment on vowel formant patterns," *J. Acoust. Soc. Am.* **109**, 748–763.

Hillenbrand, J., Getty, L. A., Clark, M. J., and Wheeler, K. (1995). "Acoustic characteristics of American English vowels," *J. Acoust. Soc. Am.* **97**, 3099–3111.

Hollien, H. (1960). "Some laryngeal correlates of vocal pitch," *J. Speech Hear. Res.* **3**, 52–58.

Huron, D. (2006). *Sweet Anticipation: Music and the Psychology of Expectation* (MIT, Cambridge, MA).

Johnston, I. (2002). *Measured Tones: The Interplay of Physics and Music* (Taylor & Francis, New York).

Johnstone, T., and Scherer, K. R. (2000). "Vocal communication of emotion," in *Handbook of Emotions*, 2nd ed., edited by M. Lewis and J. M. Haviland-Jones (Guilford, New York).

Juslin, P. N., and Laukka, P. (2003). "Communication of emotions in vocal expression and music performance: Different channels, same code?," *Psychol. Bull.* **129**, 770–814.

Krumhansl, C. L. (1990). *Cognitive Foundations for Musical Pitch* (Oxford University Press, New York).

MathWorks Inc. (2007). MATLAB (Version R2007a) (The MathWorks Inc., Natick, MA).

Nettl, B. (1956). *Music in Primitive Culture* (Harvard University Press, Cambridge, MA).

Peretz, I., Gagnon, L., and Bouchard, B. (1998). "Music and emotion: Perceptual determinants, immediacy, and isolation after brain damage," *Cognition* **68**, 111–141.

Petersen, G. E., and Barney, H. L. (1952). "Control methods used in a study of the vowels," *J. Acoust. Soc. Am.* **24**, 175–184.

Pickett, J. M. (1957). "Perception of vowels heard in noises of various spectra," *J. Acoust. Soc. Am.* **29**, 613–620.

Pierce, J. R. (1962). *The Science of Musical Sound*, revised ed. (Freeman, New York).

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, New York).

Protopapas, A., and Lieberman, P. (1997). "Fundamental frequency of phonation and perceived emotional stress," *J. Acoust. Soc. Am.* **101**, 2267–2277.

Randel, D. M. (1986). *The New Harvard Dictionary of Music*, revised 2nd ed. (Belknap, Cambridge, MA).



- Rosner, B. S., and Pickering, J. B. (1994). *Vowel Perception and Production* (Oxford University Press, New York).
- Rossing, T. D. (1990). *The Science of Sound*, 2nd ed. (Addison-Wesley, New York).
- Schellenberg, E. G., Krysciak, A. M., and Campbell, R. J. (2000). "Perceiving emotion in melody: Interactive effects of pitch and rhythm," *Music Percept.* **18**, 155–171.
- Scherer, K. R. (2003). "Vocal communication of emotion: A review of research paradigms," *Speech Commun.* **40**, 227–256.
- Scherer, K. R., Banse, R., and Wallbott, H. G. (2001). "Emotional inferences from vocal expression correlate across languages and cultures," *Journal of Cross-Cultural Psychology* **32**, 76–92.
- Schwartz, D. A., and Purves, D. (2004). "Pitch is determined by naturally occurring periodic sounds," *Hear. Res.* **194**, 31–46.
- Schwartz, D. A., Howe, C. Q., and Purves, D. (2003). "The statistical structure of human speech sounds predicts musical universals," *J. Neurosci.* **23**, 7160–7168.
- Spencer, H. (1868). *Essays: Scientific, Political, and Speculative: Volume 2* (Williams & Norgate, London).
- Stevens, K. N. (1998). *Acoustic phonetics*. (MIT, Cambridge, MA).
- Terhardt, E. (1974). "Pitch, consonance, and harmony," *J. Acoust. Soc. Am.* **55**, 1061–1069.
- Terhardt, E., Stoll, G., and Seewann, M. (1982). "Pitch of complex signals according to virtual-pitch theory: tests, examples, and predictions," *J. Acoust. Soc. Am.* **71**, 671–678.
- Thompson, W. F., and Balkwill, L. L. (2006). "Decoding speech prosody in five languages," *Semiotica* **158**, 407–424.
- Vos, P. G., and Troost, J. M. (1989). "Ascending and descending melodic intervals: Statistical findings and their perceptual relevance," *Music Percept.* **6**, 383–396.
- Zarlino, G. (1558). *Le Institutioni hamoniche*, translated by G. Marco and C. Palisca, (Yale University Press, New Haven, CT) Book 3.