

QUANTIFYING AESTHETIC PREFERENCE AND PERCEIVED COMPLEXITY FOR FRACTAL MELODIES

MICHAEL W. BEAUVOIS

Loughborough, Leics, United Kingdom

PARTICIPANTS RATED THE PERCEIVED complexity and melodiousness of fractal ($1/f^\beta$) tone sequences with specific β values. Plotting the mean complexity and melody ratings against each other and against β indicated that: (1) a melody's spectral power density slope (β) can be used as an objective measure of its perceived complexity; (2) $\beta \approx 1.50$ for optimally preferred melodies; (3) perceived complexity is determined by the distribution of pitch intervals such that optimally preferred melodies have a preponderance of small intervals compared to large ones; (4) the poor quadratic fits found in music-preference studies are due to the use of an independent-subjects design or the stimuli not covering the whole possible range of complexity; (5) $\beta < 2$ for most music stimuli; and (6) ecologically valid melodies only exist over the β range 0.67 to 2.35, with melodies whose β values lie outside this range not being perceived as musical due to their extreme complexity or simplicity. Finally, converging experimental and neurophysiological evidence is discussed that suggests that these results are a consequence of the auditory system being optimally tuned to the statistical properties of speech.

Received December 1, 2004, accepted June 8, 2006.

Key words: complexity, fractal, melody, melodiousness, preference

IT HAS BEEN CLAIMED THAT the power spectra of the pitch and loudness of successive notes in pieces of music exhibit a $1/f^\beta$ spectral density, where $\beta = 1$ (Voss & Clarke, 1975, 1978). Here, the power spectra exhibit a power-law decrease as a function of frequency that, when plotted on a log-log graph, is a straight line with a slope of $-\beta$ (see Figure 1). Similar results were reported for the pitch intervals between the successive notes of melodies (Hsu & Hsu, 1990, 1991), and these findings have been widely reported in the literature as

evidence that music is fractal in nature (e.g., Gardner, 1978; Schroeder, 1991). However, Voss & Clarke's (1975, 1978) studies have been criticised by Nettheim (1992) on the basis that their data-acquisition method was flawed, and Hsu & Hsu's (1991) study has been criticised by Henderson-Sellers & Cooper (1993) on the basis of the small range of pitch intervals examined and the fitting method used.

Furthermore, more recent studies of $1/f^\beta$ noise in music that have examined the frequency spectra of melodies have typically found that $1 < \beta < 2$. For example, Nettheim (1992) found that for melodies by Bach, Beethoven, Chopin, Mozart, and Schubert, β ranged from 1.19 to 1.88 ($M = 1.57$), and for the isochronous pitch sequences derived from these melodies, β ranged from 0.93 to 1.56 ($M = 1.33$). Similar results were found by Yadegari (1992) for the melodies of the preludes and fugues of Bach's Well-Tempered Clavier, Part I ($M = 1.47$), and by Brillinger & Irizarry (1998) for melodies taken from examples of Baroque, Classical, Romantic, Spanish guitar, Jazz, and Mambo music ($M = 1.32$, and 1.41 when Jazz and Mambo were excluded from the sample).

The tendency for music to have β values ≈ 1 has been explained by Schroeder (1991) using the results of studies in experimental aesthetics that examined the degree of arousal induced by visual stimuli. The earliest study in this area was that of Fechner (as cited in Arnheim, 1985), who systematically manipulated the proportions of the sides of rectangles and found that the rectangles most preferred by his participants were ones where the ratio of the sides accorded with the Golden Section. From this finding, and his other studies in experimental aesthetics, Fechner derived his "principle of the aesthetic middle," where experimental participants prefer stimuli that make them neither over- nor under-stimulated, but instead induce a medium degree of arousal (Arnheim, 1985). A similar conclusion was reached by Birkhoff (1933) from his studies of aesthetic preference and complexity, and resulted in his theory of aesthetic value, which stated that pleasing and interesting works of art are neither too regular and predictable, nor too irregular

and unpredictable. Schroeder (1991) used the correlation characteristics of $1/f^\beta$ tone sequences to relate Birkhoff's (1933) theory of aesthetic value to Voss & Clarke's (1975, 1978) findings, and proposed that listeners prefer music where successive notes create a melody that is moderately predictable and moderately surprising. That is, when $\beta = 2$, successive notes are highly correlated, with little note-to-note fluctuation. Hence, there will be a large proportion of small pitch intervals in the sequence, leading to a regular, predictable melody. Conversely, when $\beta = 0$, successive notes are uncorrelated, with high note-to-note fluctuation. Hence, there will be a large proportion of large pitch intervals in the sequence, leading to an irregular, unpredictable melody. According to Schroeder (1991), optimally preferred melodies occur when $\beta = 1$, where successive notes are moderately correlated with a moderate note-to-note fluctuation. Here, there will be a mixture of large and small pitch intervals in the sequence, leading to a moderately predictable and moderately surprising melody.

Schroeder's (1991) association of aesthetic preference with the β values of $1/f^\beta$ tone sequences bears some resemblance to Berlyne's (1971) theory that preference for aesthetic stimuli follows an inverted-U as a function of stimulus complexity. Here, stimuli are preferred that give rise to an optimal level of psychobiological arousal, rather than those that create very high or very low levels of arousal, similar to Fechner's "principle of the aesthetic middle." Various studies (e.g., Crozier, 1974; North & Hargreaves, 1995; Orr & Ohlsson, 2001; Vitz, 1966) have confirmed Berlyne's theory for music stimuli, finding that preference follows an inverted-U (fitted using a quadratic function) when plotted against complexity, indicating that listeners have a preferred, optimal level of stimulus variation. However, studies in this area tend to suffer from two main confounding factors. First, they often fail to quantify measures of complexity objectively, as complexity is judged subjectively by the listener. Second, the stimuli are usually excerpts from longer pieces of music, and it has been found that the liking for a piece of music is more associated with the liking of the music style, than for the piece itself (North & Hargreaves, 1997).

Schroeder's (1991) and Berlyne's (1971) theories, along with the results of the studies described above, suggest that aesthetic preference for music stimuli can be investigated using β as an objective, quantifiable measure of stimulus complexity. If this is the case, then a number of hypotheses can be made concerning the perception of $1/f^\beta$ tone sequences: (1) perceived complexity should decrease as β increases; (2) aesthetic preference

(measured by melodicty ratings) should follow an inverted-U function when plotted against both objective and perceived complexity (measured by β and complexity ratings, respectively); and (3) the β value for optimally preferred melodies (henceforth, β_{opt}) should be between 1 and 2. Some preliminary support for these hypotheses is given by the findings of two previous studies. The first, by Schmuckler & Gilden (1993; Experiment 3), investigated the sensitivity of listeners to the statistical properties of $1/f^\beta$ tone sequences by asking participants to discriminate between sequences with different β values. It was found that discrimination followed an inverted-U as a function of β , with the peak discrimination (i.e., sensitivity) occurring at $\beta \approx 2.40$. The second, by Patel & Balaban (2000), measured the amount of synchronised activity among brain regions in a 0.5-Hz band centred at 41.5 Hz in response to $1/f^\beta$ tone sequences where $\beta = 0.0, 1.30$, or 2.10 . The results indicated that the most amount of synchronised activity was generated when $\beta = 2.10$. The results of these two studies therefore not only indicate that inverted-U functions are associated with $1/f^\beta$ tone sequences, but also that listeners' responses to these sequences reach a maximum at β values > 1 , as hypothesised above.

To test the hypotheses, two experiments were conducted where participants were asked to rate the melodicty and complexity of chromatic, isochronous $1/f^\beta$ tone sequences with specific β values. Melodicty was used as a measure of music preference, as Crozier (1974) found an inverted-U relationship for ratings of melodicty against perceived complexity. Isochronous tone sequences were used on the basis of Nettheim's (1992) findings concerning the duration-spectra of melodies. Nettheim, noting that a melody is a combination of a pitch-sequence and a duration-sequence, examined the spectra of melodies encoded from Classical music scores, as well as the spectra of their individual pitch- and duration-sequences. As mentioned above, he found that β ranged from 1.19 to 1.88 ($M = 1.57$) for the melodies, and β ranged from 0.93 to 1.56 ($M = 1.33$) for the pitch-sequences. However, for the duration-sequences, β ranged from 0.0 to 0.27 ($M = 0.19$), i.e., the duration-spectra were effectively flat. Nettheim's results therefore indicate that the duration content of a melody makes a minimal contribution to its spectrum, and that the pitch content is the dominant factor. Hence, the use of isochronous sequences will: (1) remove a factor that has a minimal effect on the spectrum of a melody; (2) result in melodies with typical $1/f^\beta$ spectra; (3) simplify both the stimuli and the analysis of the results; and (4) remove rhythm and phrasing as possible confounding factors. In addition, using chromatic sequences also

avoids confounding factors such as tonality, familiarity, and music style. Furthermore, an analysis of the results with respect to the statistical properties of the sequences may also allow us to determine which characteristic of melodies determines their perceived complexity.

Experiment 1

Method

PARTICIPANTS

The participants were 18 undergraduate students (1 male, 17 female, mean age = 19.9 years, $SD = 0.8$ years), who took part in the experiment in exchange for course credit.

STIMULUS MATERIALS

Participants were asked to rate $1/f^\beta$ tone sequences with specific β values ($\beta = 0.0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8$, or 2.0) with respect to melodicty and complexity. Custom-built software written in Max 3.6.2 (Cycling74) presented the stimuli and recorded the participants' responses. The software contained four sets of unique predetermined 2,048-tone sequences for each β value investigated. Each predetermined tone sequence was generated in FORTRAN in the following manner. First, 32,768 points of a time series with a specific β value were generated using an inverse-FFT method (Timmer & König, 1995). These points then were standardised to have a mean of zero and a standard deviation of one. The standardised sequence (SS) was then scaled exponentially above and below MIDI-pitch 60 (middle C) to give a chromatic (i.e., atonal) sequence of MIDI pitches (MP):

$$\text{if } SS_i \geq 0, MP_i = \text{int} [60 * 2^{(SS_i/4)}]$$

$$\text{if } SS_i < 0, MP_i = \text{int} [60 - (60 * 2^{-(SS_i/4)} - 60)]$$

where int gives the integer part of the argument. To ensure no periodicities were present in the stimuli, only the first 2,048 points of MP were used as the predetermined tone sequence (Saupe, 1988). In comparison to other scaling methods examined by the author, the exponential scaling used here gave the best results in terms of preserving the original β value of the SS sequence when the predetermined sequences were re-analysed before use to find their β values from their spectral slopes. For all the predetermined sequences used, β was originally found to be within ± 0.001 of the desired value. However, when the sequences were

TABLE 1. Mean β Values, confidence intervals, and pitch ranges for Experiment 1 predetermined sequences.

Mean β	Mean $\pm 95\%$ CI for β	Mean pitch range (in semitones)
0.04	0.08	74.50
0.20	0.08	75.75
0.37	0.08	78.50
0.59	0.08	74.25
0.79	0.08	74.25
0.96	0.08	68.75
1.15	0.08	65.75
1.35	0.08	63.50
1.58	0.08	56.75
1.74	0.08	56.50
1.82	0.08	49.00

re-analysed after the experiment, it was found that the actual β values were lower than expected. This was due to a programming error whereby the original β values were taken from the power spectra of MP (i.e., the 32,768-point MIDI-pitch sequences), and not the first 2,048 points of MP used as the predetermined tone sequence. Here, taking subsets of points from larger sets with specific β values will result in the aliasing of high frequencies and, in effect, will add white noise to the spectra of the original $1/f^\beta$ tone sequences (Eke, Herman, Kocsis, & Kozak, 2002) resulting in lower β values than expected. However, the mean β values of the predetermined sequences were close to the originals ($\beta = 0.04, 0.20, 0.37, 0.59, 0.79, 0.96, 1.15, 1.35, 1.58, 1.74$, and 1.82) and these were noted for later use in the analysis of the results. These mean β values, the $\pm 95\%$ confidence intervals for the β values, and the mean pitch range (in semitones) of the predetermined sequences are shown in Table 1. Figure 1 shows the power spectra of five 2,048-point tone sequences created using the procedure described above, plotted on log-log graphs, along with the best linear fits to the data and the 95% confidence intervals for the β values derived from the linear fits. Figure 2 shows the first 64 notes of each of the five sequences to show the reader how different values of β result in different types of melodies. As can be seen, at low β values the melodies are angular and unpredictable, with large intervals between successive notes; at high β values the melodies are smooth and predictable, with small intervals between successive notes.

In the experiment itself, the actual stimuli presented to the participants were 128-tone MIDI-pitch sequences (MP₁₂₈) whose starting points were chosen at random for each trial from within one of the four predetermined sequences with a designated β value. In each

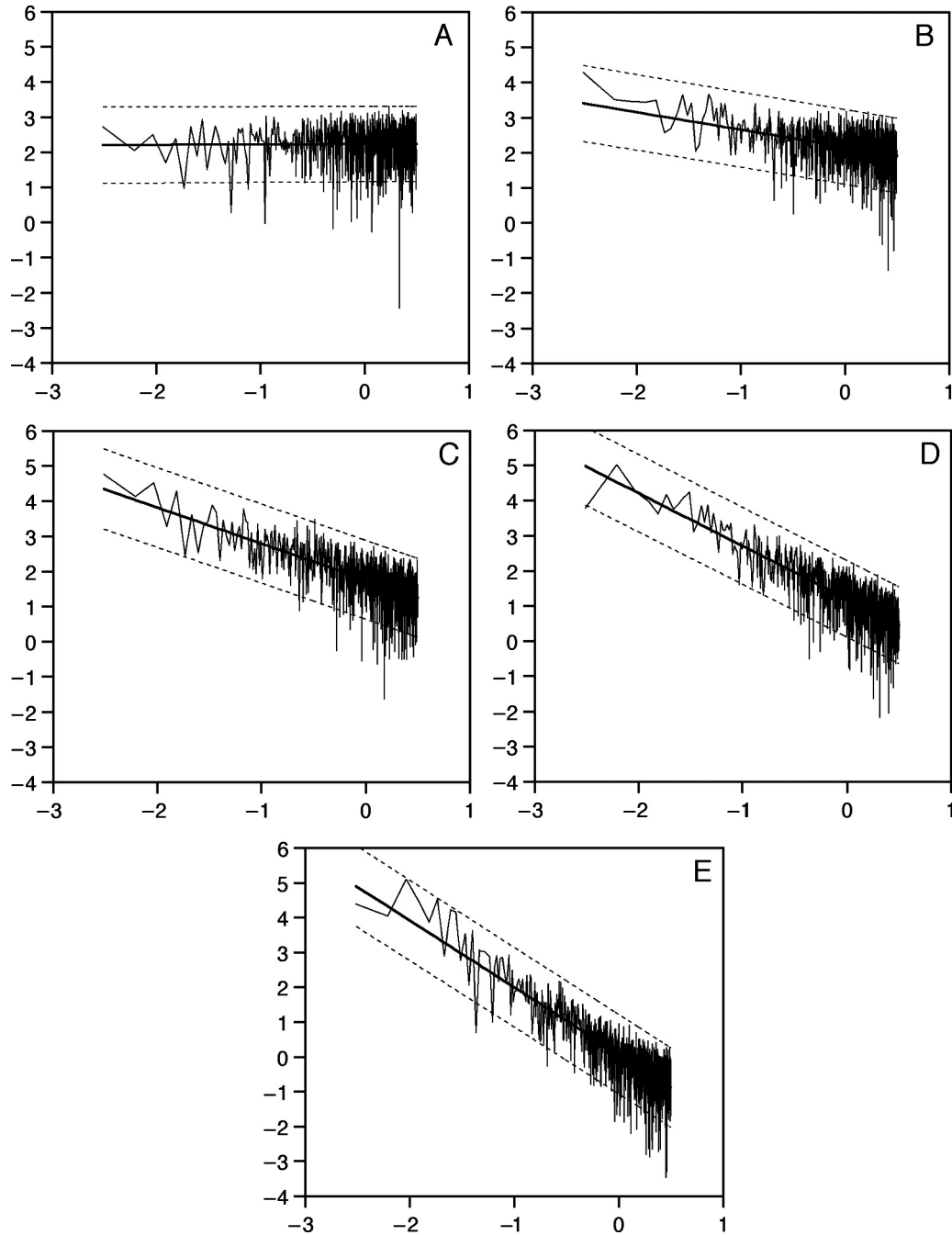


FIGURE 1. Power spectra of 2,048-point, chromatic isochronous $1/f^\beta$ pitch sequences. X-axes = log normalized frequency (radians/sample), y-axes = log power (power/radians/sample). Also shown are the best linear fits to the data (bold lines) and the 95% confidence intervals (CI) for the β values derived from the linear fits. For Figure 1a, $\beta = -0.01$, 95% CI = ± 0.08 ; for Figure 1b, $\beta = 0.49$, 95% CI = ± 0.08 ; for Figure 1c, $\beta = 1.03$, 95% CI = ± 0.08 ; for Figure 1d, $\beta = 1.5$, 95% CI = ± 0.08 ; for Figure 1e, $\beta = 1.92$, 95% CI = ± 0.08 .

MP_{128} sequence, the onset-to-onset time and duration of the tones were 250 milliseconds, and each tone had a MIDI velocity of 80. Each MP_{128} sequence lasted 32 seconds and triggered a Yamaha YS-100 Digital Synthesiser set to preset 40 "Piano 2" with no reverberation. The

synthesiser's audio output went into a Yamaha AX-592 amplifier, and the amplifier's audio output was fed into a soundproof cabin and was presented dichotically to participants over Sennheiser HD 520-II headphones at 85 dB SPL (measured with a flat-plate

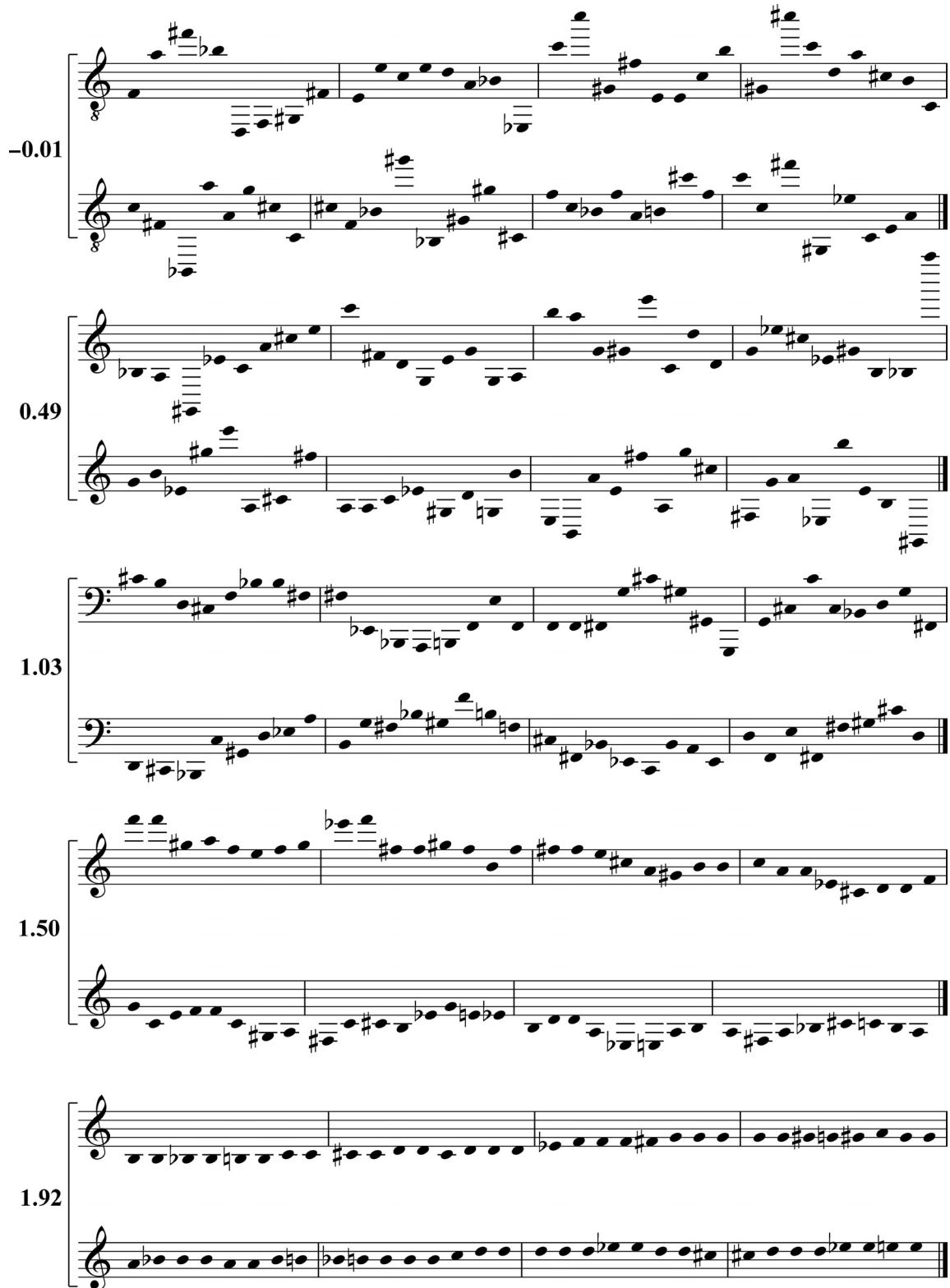


FIGURE 2. The first 64 notes of the isochronous $1/f^\beta$ pitch sequences whose spectra are shown in Figure 1, with bars added for ease of reading. Accidentals remain in force within each bar, unless cancelled. For each sequence, the β value is shown to the left of the relevant staves.

coupler). At the start of the experiment, and after the participant made their response, there was a 5-second silence before the next stimulus was presented.

PROCEDURE

The participants sat in a soundproof cabin wearing headphones and could see a computer screen through a built-in glass panel. The screen was connected to a Power Macintosh 7200/75 and displayed an interface to the custom-built software. The cabin also contained a keyboard and a mouse that the participants used to make their responses.

The software interface for the complexity condition displayed rating buttons numbered 0 to 10 from left to right, with 0 labeled "least complex," 5 labeled "mid-way," and 10 labeled "most complex." The interface instructed the participants to: (1) start a new block and listen to the tone sequence; (2) rate the complexity of the sequence; and (3) press the appropriate rating button, then listen to the next sequence. For the melody condition, the interface was exactly the same, except that the word "complex" was replaced by "melodic."

Before each condition, participants read a sheet of paper containing task instructions for either the complexity or melody conditions. The instructions for both conditions followed the same format and informed the participants that: (a) they would hear 11 tone sequences, each sequence lasting for 32 seconds; (b) their task was to rate how complex or melodic they thought the sequences were, independent of the level of complexity or melodiousness in the music they normally listened to; and (c) they should try to use the full range of the rating scale. The instructions also defined 'complex' as being how easy it was to predict what the sequence would do next and how many surprises the sequence contained, and 'melodic' as being how melodious, tuneful, and musical the sequence was. The extremes of the rating scales also were defined, with 0 being the least complex (i.e., the most predictable, simple, and uniform) or the least melodic (i.e., the most unmelodious, tuneless, and unmusical), and 10 being the most complex (i.e., the most unpredictable, surprising, and erratic) or the most melodic (i.e., the most melodious, tuneful, and musical). Ratings of 5 were defined as being midway between the two extremes.

For both conditions, participants read the instructions, completed one practice block, and then 10 blocks with a short break between blocks 5 and 6. Each block presented one MP₁₂₈ sequence for each of the eleven β values in random order. Participants completed both the melody and complexity conditions, and the order of

the conditions was counterbalanced across participants. The experiment took approximately 2.5 hours to complete.

Results & Discussion

CALCULATION OF β_{OPT}

Two graphs were plotted of the mean melody and complexity ratings across blocks 1 to 10. The first graph (Figure 3a) shows melody (*Mel*) plotted against complexity (*Comp*). Hypothesis (2) of the Introduction stated that melody should follow an inverted-U (quadratic) function when plotted against perceived complexity, and Figure 3a appears to confirm this. The presence of quadratic trends was verified by fitting the data with exponential, linear and quadratic functions and noting the adjusted R^2 values for each fit. The best fit to the data (shown in Figure 3a) was given by a quadratic function (adjusted $R^2 = .92$), thereby confirming Hypothesis (2).

The second graph (Figure 3b) shows the mean melody and complexity ratings plotted against β . Figure 3b indicates that perceived complexity decreases as β increases, and that melody follows a curvilinear function against objective complexity (β) that reaches a peak at high β values, thereby confirming Hypotheses (1) and (2) of the Introduction. However, Figure 3b also shows clear asymptotes for both the complexity and melody ratings when $\beta < 0.40$ to 0.60 . This results in a linear rise in complexity as β decreases that reaches an asymptote at low β values after a curved crossover region between the asymptote and the ascending linear trend, and a peak for melody at high β values that flattens to an asymptote at low β values after a curved crossover region between the asymptote and the curvilinear trend. These asymptotes for complexity and melody are also present in Figure 3a, where they cause a bunching of data points on the right-hand side of the graph (i.e., for low β /high-complexity stimuli).

To define the underlying response trends in the data and quantify the value of β_{OPT} the complexity and melody responses shown in Figure 3b were fitted with a collection of functions (exponential, linear, quadratic, 4-parameter sigmoidal, and 4-parameter Gaussian), and the best fit was taken to be the one with the highest adjusted R^2 value. To obtain meaningful parameter values for the sigmoidal and Gaussian fits, a constrained fitting procedure was used where the upper and lower asymptotes of the sigmoidal function and the peak and baseline of the Gaussian function fell between the highest and lowest possible scores for complexity

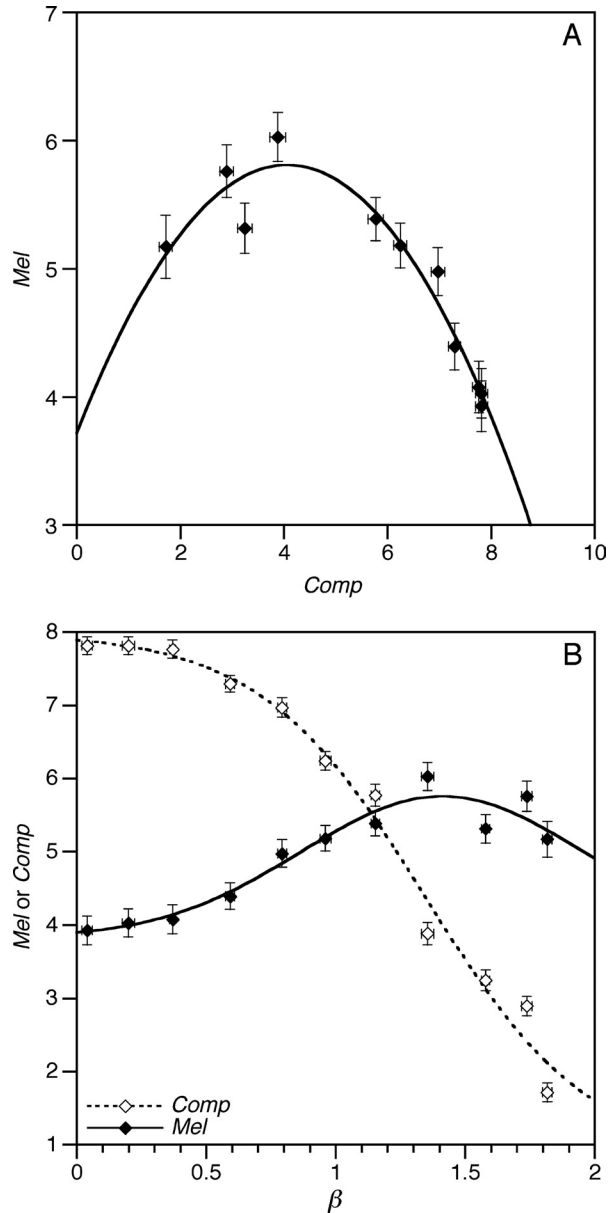


FIGURE 3. (A) Mean melody ratings (*Mel*) plotted against mean complexity ratings (*Comp*) from Experiment 1. Also shown is the quadratic best-fit to the data (solid line): $Mel = -0.13 \text{ Comp}^2 + 1.03 \text{ Comp} + 3.72$ [$R^2 = .94$, adjusted $R^2 = .92$]. Standard-error bars also are shown for the values of *Comp* ($N = 18$), and *Mel* ($N = 18$). b) Mean melody (*Mel*) and complexity (*Comp*) ratings from Experiment 1 plotted against β . Also shown are the sigmoidal and Gaussian best-fits to the data. Sigmoidal fit (dotted line) to mean complexity ratings (*Comp*): $Comp = 7.99 + (0.79 - 7.99) / (1 + \exp((1.34 - \beta) * 3.14))$ [$R^2 = .98$, adjusted $R^2 = .98$]. Gaussian fit (solid line) to mean melody ratings (*Mel*): $Mel = 1.93 \exp(-0.5((\beta - 1.41)/0.55)^2) + 3.83$ [$R^2 = .93$, adjusted $R^2 = .90$]. Standard-error bars also are shown for the values of β ($N = 4$), *Comp* ($N = 18$), and *Mel* ($N = 18$).

and melody (i.e., 0 and 10). It was found that the best fit to the complexity data was given by the sigmoidal function (adjusted $R^2 = .98$) and the best fit to the melody data was given by the Gaussian function (adjusted $R^2 = .90$). Both these fits are shown in Figure 3b. Values for β_{opt} can be found from the midpoint of the sigmoidal fit and the peak value (i.e., the highest melody rating) of the Gaussian fit. These occurred at $\beta = 1.34$ for the sigmoidal fit, and at $\beta = 1.41$ for the Gaussian fit, giving a mean β_{opt} of 1.38, thereby confirming Hypothesis (3) of the Introduction.

The melody and complexity ratings shown in Figure 3b also allow us to verify the value of $\beta_{opt} = 1.38$ found above. First, the complexity value corresponding to the peak value of the quadratic function shown in Figure 3a was calculated. This complexity value was then entered into the formula for the sigmoidal fit to the complexity ratings in Figure 3b to find the β value corresponding to the quadratic-function peak in Figure 3a. This gave a β value of 1.40, which not only falls within the β_{opt} range found above ($\beta = 1.34$ to 1.41), but is also close to the mean β_{opt} value found above ($\beta_{opt} = 1.38$).

INTERVAL DISTRIBUTIONS AND COMPLEXITY

The results reported above suggest that the perceived complexity of melodies may be due to the distribution characteristics of the pitch intervals between successive notes. For example, for a $1/f^\beta$ tone sequence where $\beta = 0$, the equal probability of large and small intervals occurring results in a large standard deviation of pitch-interval size (interval-SD) around a mean of zero and, consequently, a high complexity rating. However, as β increases, the probability of large intervals occurring decreases until, when $\beta = 2$, the sequence is composed mainly of small intervals, leading to a small interval-SD and a low complexity rating. To confirm the association of perceived complexity with β and interval-SD, two-tailed correlations were carried out between β , the mean complexity ratings given by the participants (*Comp*), and the mean interval-SDs of the predetermined tone sequences used in Experiment 1. All correlations were highly significant. β was negatively correlated with interval-SD, $r(9) = -1.00$, $p < .001$, and *Comp*, $r(9) = -.96$, $p < .001$, indicating that both interval-SD and perceived complexity increase as β decreases. This finding also is reflected in the positive correlation of interval-SD with *Comp*, $r(9) = .96$, $p < .001$. These findings therefore confirm the association of perceived complexity with β , and support the hypothesis that the distribution of pitch intervals in a melody determines its perceived complexity.

Experiment 2

The results of Experiment 1 confirmed the three hypotheses made in the Introduction, and indicated that the distribution of pitch intervals determines the perceived complexity of a melody. However, it could be argued that these results are flawed due to the stimuli being created by taking subsets of points from larger sets with specific β values. As described above, this procedure will result in the aliasing of high frequencies and, in effect, will add white noise to the spectra of the original $1/f^\beta$ tone sequences (Eke et al., 2002). Consequently, the β values of the subsets used as the stimuli in Experiment 1 will be lower than expected, and the actual β_{opt} will be < 1.38 . In addition, experiments investigating music preference and complexity (e.g., North & Hargreaves, 1995; Orr & Ohlsson, 2001) typically use an independent-subjects design where participants rate the stimuli only for complexity *or* preference in order to eliminate any artifacts in the data arising from the participants' own hypotheses about the stimulus variables. Therefore, to address these issues and confirm the findings of Experiment 1, the participants in Experiment 2 rated individual sequences with specific β values for either melodicty *or* complexity.

Method

PARTICIPANTS

The participants were 86 undergraduate students (20 male, 66 female, mean age = 19.2 years, $SD = 3.0$ years) who took part in the experiment in exchange for course credit.

STIMULI

Participants were asked to rate $1/f^\beta$ tone sequences with specific β values ($\beta = 0.0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0, 2.2$, or 2.4) with respect to melodicty *or* complexity. The tone sequences were generated in exactly the same manner as in Experiment 1, except that only the first 128 points of MP were used to create a predetermined tone sequence with a specific β value. When the predetermined sequences were re-analysed to find their β values from their spectral slopes, it was found that the actual β values were lower than expected. This effect was probably due to high-frequency aliasing (as described in Experiment 1 above and Eke et al., 2002). Therefore, the actual β value of each sequence was noted for later use in the analysis of the results.

As in Experiment 1, the stimuli were chromatic 128-tone MIDI-pitch sequences (MP_{128}) where the onset-to-onset

TABLE 2. Mean β values, confidence intervals, and pitch ranges for Experiment 2 128-tone MIDI-Pitch sequences (MP_{128}).

Mean β	Mean $\pm 95\%$ CI for β	Mean pitch range (in semitones)
-0.73	0.37	67.92
-0.49	0.36	67.52
-0.26	0.34	70.82
-0.03	0.37	68.35
0.19	0.35	68.00
0.41	0.35	68.10
0.63	0.35	66.82
0.84	0.35	64.82
1.05	0.35	66.37
1.26	0.36	63.02
1.45	0.34	60.52
1.66	0.23	51.52
1.84	0.18	43.55

time and duration of the tones in each MP_{128} sequence was 250 milliseconds, and each tone had a MIDI velocity of 80. Each MP_{128} sequence lasted 32 seconds and was unique. The mean β values, the $\pm 95\%$ confidence intervals for the β values, and the mean pitch range (in semitones) of the MP_{128} sequences are shown in Table 2. The MP_{128} sequences were converted to MIDI files that then were converted to AIFF files using Virtual Sound Canvas 3 (Roland) software set to preset 0 "Piano 1" with no reverberation. The AIFF files of the stimuli were burnt onto four pairs of CDs, each CD-pair containing 10 examples of the 13 β values in a different random order. A Yamaha CDX-390 CD player presented the stimuli via a Yamaha AX-592 amplifier and B&W bookshelf loudspeakers placed approximately five feet apart at 90 dB SPL (measured one foot from the loudspeaker).

PROCEDURE

Participants were tested in groups of three to seven at a time in a quiet room, seated in an inverted-U arrangement facing the loudspeakers. This procedure served to minimise the possibility of distraction within the group (North & Hargreaves, 1995). Unlike Experiment 1, participants in each experimental session were assigned to either the complexity *or* the melodicty condition. For both conditions, participants read the instructions (which were the same as Experiment 1), completed one practice block (which contained one example of each of the 13 stimuli), and then listened to one of the stimulus CD-pairs. Participants rated on

answer sheets how complex or melodic they thought each stimulus sequence was. The rating scales were 11-point scales ranging from 0 (least melodic/complex) to 10 (most melodic/complex), as in Experiment 1. There was an 8-second gap between sequences for the participants to make their responses. The use of the stimulus CD-pairs was counterbalanced across participants, and the experiment took approximately 1.5 hours to complete. At the end of the experiment, participants were given a sheet of paper where they were asked to rate the degree of attention they paid to the stimuli on an 11-point scale, where 5 was a point midway between 0 (complete lack of attention) and 10 (total attention).

Results & Discussion

Calculation of β_{opt}

Product-moment correlations (Pearson's r , two-tailed) were carried out between the stimulus CD-pairs for the complexity and melodicity ratings to test for any possible order effects in the stimuli. All correlations between CD-pairs for the complexity ratings were significant, $r(11) \geq .90, p < .001$. However, for the melodicity ratings, all correlations between the CD-pairs were significant, $r(11) \geq .65, p \leq .02$, except for one CD-pair that did not significantly correlate with any of the other pairs, $r(11) \leq .42, p \geq .15$. Therefore, the data obtained from this CD-pair were excluded from further analysis and the complexity and melodicity data were pooled across the remaining three CD-pairs. In addition, participants who gave ratings of < 5 for the attention they paid to the stimuli were excluded from the analysis. The final results, therefore, were obtained from 25 participants for the complexity condition, and 23 participants for the melodicity condition.

A preliminary investigation of the participants' responses using cluster analysis (Ward's method, squared Euclidean distance) indicated the presence of two distinct participant groups in both the complexity and melodicity conditions. One group in the complexity condition (CG2) exhibited a monotonic decline in complexity ratings with increasing β , and one group in the melodicity condition (MG2) exhibited a curvilinear function for melodicity ratings that reached a peak at high β values. For both these groups, there were clear asymptotes for the complexity and melodicity ratings when $\beta < 0.40$ to 0.60 , and both patterns of responses agreed with those found in Experiment 1 for complexity and melodicity—(see Figure 4b). However, the other participant groups in each condition gave uniform ratings over the whole

range of β values investigated, resulting in linear response patterns for the complexity (CG1) and the melodicity (MG1) conditions (see Figure 4b). The presence of these CG1 and MG1 groups may be due to the chromatic/atonal nature of the stimuli employed in the experiment. For example, the CG1 participants may have confused atonality with complexity and, consequently, rated all the stimuli as being equally complex, leading to uniformly high complexity ratings over the whole range of β values. Similarly, the MG1 participants may have confused atonality with unmelodiousness and rated all the stimuli as being equally unmelodic, leading to uniformly low melodicity ratings over the whole range of β values. Alternatively, the presence of the CG1 and MG1 groups may be due to the range of β values used in Experiment 2. In Experiment 1, β ranged from 0.04 to 1.82 . However, in Experiment 2, β ranged from -0.73 to 1.84 . This suggests that the predominance of stimuli with low β values in Experiment 2 (i.e., high-complexity/low-melodicity stimuli) may have biased some participants towards giving high complexity and low melodicity ratings across the whole β range. Further examination of the participant groups also revealed the following gender differences: for CG1, $N = 8$ (2 male, 6 female), for CG2, $N = 17$ (4 male, 13 female), for MG1, $N = 11$ (5 male, 6 female), and for MG2, $N = 12$ (0 male, 12 female). This initially would appear to suggest a gender difference in the perception of melodicity in $1/f^\beta$ tone sequences. However, this is unlikely to be the case for three reasons. Firstly, a Kruskal-Wallis one-way ANOVA performed on the four cluster groups resulted in $\chi^2(3) = 6.61$ and $p = .08$ for gender, indicating no significant gender differences between cluster groups. Secondly, a cluster analysis (Ward's method, squared Euclidean distance) of the participants' combined complexity and melodicity responses for Experiment 1 showed no obvious participant groups. Finally, the responses of the sole male participant in Experiment 1 followed *exactly* the same trends as the female participants for complexity and melodicity. These results therefore suggest that confusion or response bias was responsible for the CG1 and MG1 responses, and that the greater number of males in MG1 occurred simply by chance and seemed to appear significant because of the relatively small sample size. However, regardless of the reasons for the presence of the CG1 and MG1 groups, their data clearly were not applicable, and were excluded from further analysis.

These participant groups being excluded, the same procedure was used as in Experiment 1 to analyse the data. Two graphs were plotted of the mean complexity

and melodicity ratings for all participant groups over the range $\beta = -0.03$ to 1.84. The data points for $\beta < -0.03$ were excluded to enable the best comparison with the results of Experiment 1, and also because no additional information was given by these data points as the asymptotes resulted in level traces for complexity and melodicity below this point. The first graph (Figure 4a) shows the melodicity responses plotted against the complexity responses and indicates that, for CG2 and MG2, melodicity follows a quadratic trend against complexity, as found in Experiment 1 (see Figure 3a). The quadratic trend was verified by fitting the data with exponential, linear, and quadratic functions, and noting the adjusted R^2 values for each fit. The best fit to the data was given by a quadratic function (adjusted $R^2 = .94$), confirming the quadratic trend for melodicity against complexity found in Experiment 1.

The second graph (Figure 4b) shows the mean complexity and melodicity ratings for all participant groups plotted against β . Figure 4b clearly shows that the responses of the CG1 and the MG1 groups are uniform over the whole β range, and that the CG2 and MG2 groups exhibit the same trends for complexity and melodicity shown by the participants in Experiment 1. However, the overall rating range is smaller compared to Experiment 1, and this is probably due to the differences between Experiments 1 and 2. In Experiment 1, each participant was tested alone in a soundproof cabin and rated the stimuli for complexity *and* melodicity. In Experiment 2, each participant was tested in a group in a quiet room and rated the stimuli for complexity *or* melodicity. Consequently, participants in Experiment 1, who would have had more exposure to the full range of stimulus complexity and melodicity compared to the participants in Experiment 2, might have been better able to discriminate between the different $1/f^\beta$ sequences for the second condition they completed. As the complexity and melodicity conditions in Experiment 1 were counterbalanced across participants, this may have resulted in the wider rating range found in Experiment 1.

Next, to define the underlying response trends in the data and quantify the value of β_{opt} , the complexity and melodicity responses were fitted with the same collection of functions used in Experiment 1, using the same constrained fitting procedure for the sigmoidal and Gaussian fits, and the best fit was taken to be the one with the highest adjusted R^2 value. The best fit to the CG2 data was given by the sigmoidal function (adjusted $R^2 = .96$), and the best fit to the MG2 data

was given by the Gaussian function (adjusted $R^2 = .90$). Both these fits are shown in Figure 4b. As in Experiment 1, values for β_{opt} can be found from the midpoint of the sigmoidal fit and the peak value of the Gaussian fit. These occurred at $\beta = 1.68$ for the sigmoidal fit, and at $\beta = 1.60$ for the Gaussian fit, giving a mean β_{opt} of 1.64 for Experiment 2. This β_{opt} value is fairly close to the mean β_{opt} of 1.38 found in Experiment 1, and suggests that the same perceptual effects were being measured by both experiments.

As in Experiment 1, the melodicity and complexity ratings shown in Figure 4b allow us to verify the values of β_{opt} found above. First, the complexity value corresponding to the peak value of the quadratic function in Figure 4a was calculated. This complexity value was then entered into the formula for the sigmoidal fit to the complexity ratings in Figure 4b to find the β value corresponding to the quadratic-function peak in Figure 4a. This gave a β value of 1.62, which not only falls within the β_{opt} range found above ($\beta = 1.60$ to 1.68), but also is close to the mean β_{opt} value of 1.64 found above. Furthermore, as the mean β_{opt} found in Experiment 1 was 1.38, the mean β_{opt} of 1.64 found in Experiment 2 suggests an overall value of $\beta_{opt} \approx 1.50$, a value that agrees well with the mean β values found by Nettheim (1992) for classical melodies (1.57), by Ydegari (1992) for the prelude and fugue melodies of Bach's Well-Tempered Clavier Part I (1.47), and by Brillinger & Irizarry (1998) for melodies taken from examples of Baroque, Classical, Romantic, and Spanish guitar music (1.41).

INTERVAL DISTRIBUTIONS AND COMPLEXITY

To confirm the association between perceived complexity (*Comp*), β , and interval-SD found in Experiment 1, two-tailed correlations were carried out between the β values, the complexity ratings given by CG2, and the mean interval-SDs of the tone sequences used in Experiment 2. All data points, including those for $\beta < -0.03$, were used in this analysis, as the asymptotes described above may have been associated with interval-SD. All correlations were highly significant, and the results paralleled those of Experiment 1. That is, β was negatively correlated with interval-SD, $r(11) = -.98$, $p < .001$, and *Comp*, $r(11) = -.79$, $p = .001$, indicating that both interval-SD and perceived complexity increase as β decreases. Interval-SD was also positively correlated with *Comp*, $r(11) = .89$, $p < .001$. These results, along with the similar values found for β_{opt} , therefore confirm the findings of Experiment 1 and the hypotheses made in the Introduction.

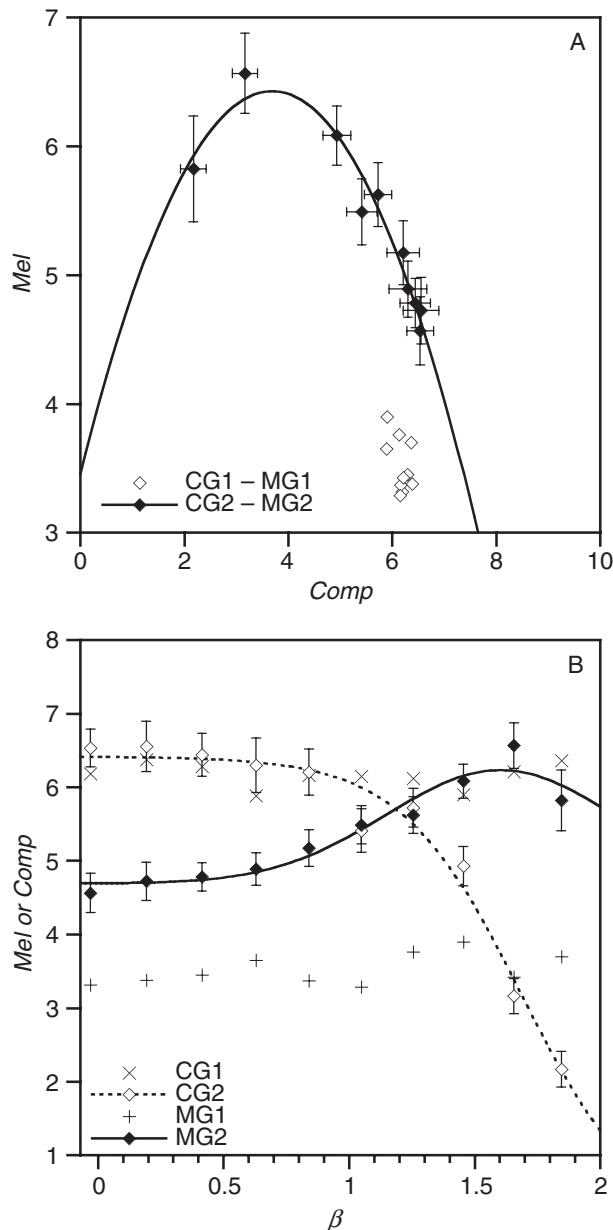


FIGURE 4. (A) Mean melody ratings (*Mel*) plotted against mean complexity ratings (*Comp*) from Experiment 2 for participant groups CG1 and MG1, and CG2 and MG2. Also shown is the quadratic best-fit to the data for CG2 and MG2 (solid line): $Mel = -0.22 Comp^2 + 1.61 Comp + 3.46$ [$R^2 = .95$, adjusted $R^2 = .94$]. Standard-error bars also are shown for the values of *Comp* ($N = 17$), and *Mel* ($N = 12$). (B) Mean melody ratings (*Mel*) and complexity (*Comp*) ratings from Experiment 2 plotted against β for participant groups CG1, CG2, MG1, and MG2. Also shown are the sigmoidal and Gaussian best-fits to the data for CG2 and MG2. Sigmoidal fit (dotted line) to mean complexity ratings (*Comp*): $Comp = 6.42 + (0 - 6.42) / (1 + \exp((1.68 - \beta) * 4.2))$ [$R^2 = .97$, adjusted $R^2 = .96$]. Gaussian fit (solid line) to mean melody ratings (*Mel*): $Mel = 1.54 \exp(-0.5((\beta - 1.6)/0.45)^2) + 4.69$ [$R^2 = .94$, adjusted $R^2 = .90$]. Standard-error bars also are shown for the values of β ($N = 30$), *Comp* ($N = 17$), and *Mel* ($N = 12$).

Quadratic Fitting and Ecological Validity

A number of studies that have investigated music preference as a function of complexity have found that, although the best fits to the data are given by quadratic functions, the fits typically are sparse on the low-complexity side of the functions (North & Hargreaves, 1995; Orr & Ohlsson, 2001). The same effect was found in this study when the melody ratings were plotted against the corresponding complexity ratings (see Figures 3a and 4a). Here, the best fits to the data from Experiments 1 and 2 were quadratic functions ($R^2 = .94$ and $.95$, respectively) with sparse fits on the low-complexity (i.e., high β) side of the functions. The association of these sparse-fit areas with high β values suggests that this effect has occurred in previous studies because of two factors associated with the design considerations of the experiments.

The first factor is the nature of the inverted-U function of aesthetic preference against complexity. According to Berlyne (1971), the inverted-U function arises as the result of the combined action of two biological systems that exhibit opposing sigmoidal trends with respect to affect as stimulus complexity (or arousal) increases. The first, a reward/pleasure system, generates a positive affect; the second, an aversion/displeasure system, generates a negative affect. The activation of both systems increases with stimulus complexity, such that affective asymptotes are reached at high levels of complexity. Here, the combination (or sum) of the two opposing sigmoidal reward and aversion functions results in a curvilinear, inverted-U function that determines the hedonic value of the stimuli, with participants preferring stimuli of medium complexity where the hedonic value is maximal. However, the results of Experiments 1 and 2 suggest that the inverted-U function can instead be considered as being due to the interaction between sigmoidal and Gaussian functions for complexity and melody/preference. This interaction, and the nonlinear nature of the sigmoidal and Gaussian functions, suggests a possible explanation for the sparse-fit areas, which can be illustrated in the following manner. First, the sigmoidal and Gaussian fits to the complexity and melody data in Experiments 1 and 2 are used to create complexity-melody (CompMel) functions equivalent to the trends underlying the participant responses shown in Figures 3a and 4a. That is, the Gaussian fits to the participants' mean melody ratings (*Mel*) are plotted against the sigmoidal fits to the participants' mean complexity ratings (*Comp*) for

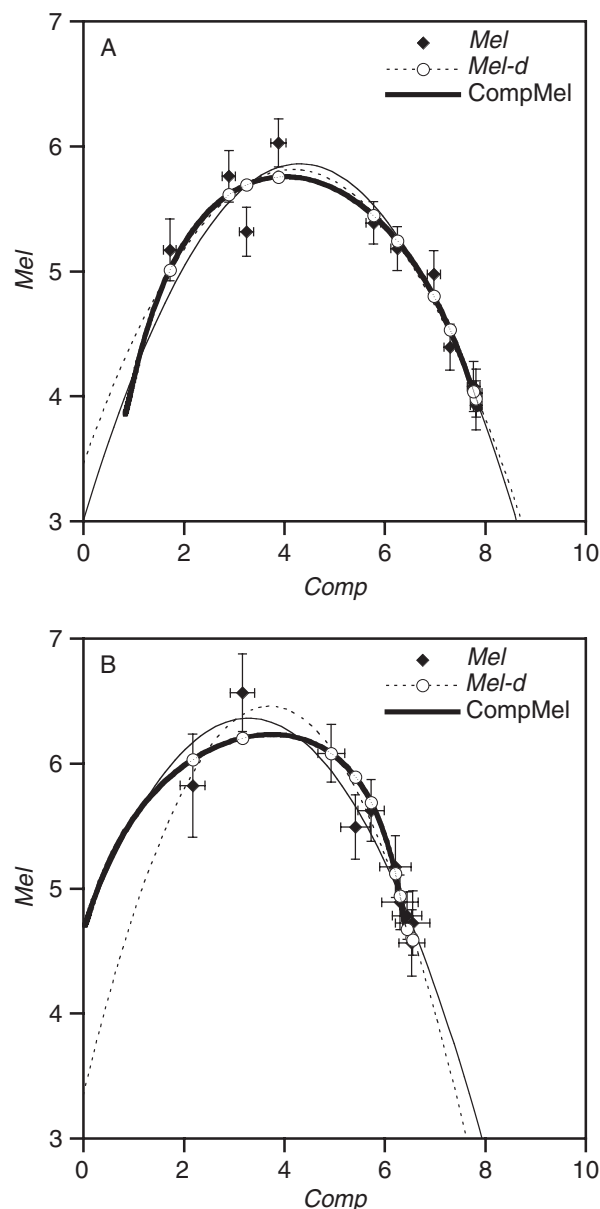


FIGURE 5. (A) Complexity-melodicity (CompMel) function (thick line) created from sigmoidal and Gaussian fits to the complexity and melodicity data in Experiment 1. Also shown are the melodicity ratings given by participants (*Mel*) from Figure 3a and those derived from the CompMel function (*Mel-d*). Quadratic fits: to CompMel function (solid line), $Mel = -0.15 \text{ Comp}^2 + 1.32 \text{ Comp} + 3.0$ [$R^2 = .98$]; to *Mel-d* (dotted line), $Mel = -0.13 \text{ Comp}^2 + 1.13 \text{ Comp} + 3.46$ [$R^2 = 1$]. b) Complexity-melodicity (CompMel) function (thick line) created from sigmoidal and Gaussian fits to the complexity and melodicity data in Experiment 2. Also shown are the melodicity ratings given by participants (*Mel*) from Figure 4a and those derived from the CompMel function (*Mel-d*). Quadratic fits: to CompMel function (solid line), $Mel = -0.16 \text{ Comp}^2 + 1.02 \text{ Comp} + 4.68$ [$R^2 = .95$]; to *Mel-d* (dotted line), $Mel = -0.23 \text{ Comp}^2 + 1.68 \text{ Comp} + 3.35$ [$R^2 = .97$].

both experiments over the range $0 \leq \beta \leq 3$. To simulate the participants' responses to the stimuli, the *Comp* values from each experiment are used to derive the equivalent melodicity values (*Mel-d*) from the CompMel functions. For each experiment, *Mel* and *Mel-d* are then plotted against *Comp*, along with the CompMel functions themselves (see Figure 5), and the CompMel functions and *Mel-d* fitted with quadratics. For Experiment 1, the CompMel function and *Mel-d* are well fitted ($R^2 = .98$ and 1, respectively; see Figure 5a). However, for Experiment 2, the fits to the CompMel function and *Mel-d* are less good ($R^2 = .95$ and .97, respectively; see Figure 5b). The reason for this can be seen in Figure 5, which shows that the CompMel function and the *Mel-d* values for Experiment 2 follow a distorted quadratic trend that results in a worse quadratic fit and a sparser fit to low-complexity data compared to Experiment 1. The closer similarity of the results of Experiment 2 to those of North & Hargreaves (1995) and Orr & Ohlsson (2001) suggests that the sparse fits to low-complexity data are due to the independent-subjects design employed by all three studies. This conclusion is supported by the results for Experiment 1 (shown in Figure 5a), that show a more symmetrical and quadratic CompMel function, a less sparse fit to low-complexity data, and a larger response range compared to Experiment 2. These results also suggest that: (1) although inverted-U functions for preference against complexity are well approximated by quadratics, caution should be exercised when analysing/fitting the data gathered using an independent-subjects design; and (2) where possible, aesthetic-preference experiments should use a procedure and stimuli similar to Experiment 1 to obtain clearer response trends.

The second factor associated with the design considerations of previous studies that may account for the sparse-fit areas is the ecological validity of the stimuli. For the stimuli in music-preference studies to have ecological validity (i.e., be musical), the stimuli cannot fall over the whole possible range of complexity. This occurs because most music stimuli exhibit some degree of change, and simpler stimuli might not be perceived as music as such by listeners. Therefore, extremely simple stimuli such as the monotonous repeated-note sequences one obtains when $\beta > 2$ have been excluded from previous studies for being "unmusical." This is supported by Orr & Ohlsson's (2001) statement regarding the musicians who created the stimuli used in their study: "... when instructed to produce something very simple, they nevertheless played music. They did not,

for example, play the same tone over and over again” (p. 123). As the sparse-fitting effect for low-complexity also occurs for music stimuli that are excerpts from pre-existing pieces (North & Hargreaves, 1995), the replication of this effect using the data from Experiments 1 and 2 therefore: 1) supports Orr & Ohlsson’s (2001) suggestion that “the lower bound on complexity . . . might reflect the nature of music rather than a methodological difficulty” (p. 124); and 2) indicates that $\beta < 2$ for most music stimuli.

Our replication of the sparse quadratic fits found by North & Hargreaves (1995) and Orr & Ohlsson (2001) also suggests that the same data might allow us to estimate the range of β values over which ecologically valid melodies exist. Here, the bounds of the ecologically valid β range define the upper and lower limits of melodic complexity, beyond which melodies are not perceived as being musical due to their extreme complexity or simplicity. These bounds can be estimated by examining the asymptotes of the sigmoidal and Gaussian fits to the data at extreme β values. That is, the upper and lower bounds for ecologically valid melodies are indicated by: (1) the asymptotes of the sigmoidal fits to the complexity data; and (2) the decline of the Gaussian fits to the melodicty data to a baseline. To assess the ecologically valid range for melodies, the criteria for the upper and lower bounds were taken to be the β values where each of the sigmoidal fits to the complexity data reached 10% and 90% of their full range (i.e., upper-lower asymptote) and each of the Gaussian fits to the complexity data reached 10% of their full range (i.e., peak-baseline). For Experiment 1, these points occurred at $\beta = 0.64$ and 2.04 for complexity (*Comp*) and $\beta = 0.24$ and 2.58 for melodicty (*Mel*), resulting in mean β values of 0.44 for the lower bound and 2.31 for the upper bound. For Experiment 2, these points occurred at $\beta = 1.16$ and 2.20 for CG2 and $\beta = 0.63$ and 2.57 for MG2, resulting in mean β values of 0.89 for the lower bound and 2.39 for the upper bound. The mean upper and lower bound values taken from these results therefore indicate an ecologically valid β range of 0.67 to 2.35 to accompany the overall β_{opt} of ≈ 1.50 found above. Support for the validity of this ecologically valid β range is given by: (1) its convergence with the lowest and highest β values (0.73 and 1.90, respectively) for the melodies investigated by Brillinger & Irizarry (1998) and Yadegari (1992); and (2) its midpoint, which occurs at $\beta = 1.51$, a value very close to the overall β_{opt} of ≈ 1.50 derived from the results of Experiments 1 and 2.

Interval Distributions and Melodies

The results of Experiments 1 and 2 indicated that the distribution of pitch intervals in a melody determines its perceived complexity. For the overall β_{opt} of ≈ 1.50 , the interval-SD associated with this β value will result in melodies with a preponderance of small pitch intervals compared to large ones, a characteristic that applies to melodies from a wide variety of cultures (Huron, 2001; von Hippel, 2000; Vos & Troost, 1989). For example, von Hippel (2000, Figures 1 and 4) plotted the percentage occurrence of pitch intervals against absolute interval size in semitones for samples of melodies taken from Schubert lieder and folk-song collections (Chinese, European, Ojibway [Native American], and South African), and found that interval occurrence decreased as a function of absolute interval size. The same pattern of results also were found by Vos & Troost (1989, Figure 2) for samples of melodies written by composers (Bach, Bartok, Beethoven, Brahms, Chopin, Debussy, Dvorak, Mozart, Schubert, Schumann, Shostakovich, J. Strauss, and Stravinsky) and “ethnomusic” (Albanian, Bulgarian, Iberian, Irish, Macedonian, Norwegian, Sicilian, and American Negro folk songs).

As described previously, $1/f^\beta$ tone sequences have specific pitch-interval distributions that vary as a function of β . This suggests that, if the overall β_{opt} of ≈ 1.50 found from Experiments 1 and 2 is valid, von Hippel’s (2000) and Vos & Troost’s (1989) interval data should correlate highest with the interval data from $1/f^\beta$ tone sequences with this β value. This hypothesis can be confirmed in the following manner. First, the sets of four 2,048-point predetermined sequences used in Experiment 1 to generate $1/f^\beta$ tone sequences with specific β values were examined, and the mean percentage occurrence of each absolute interval size over each set of four sequences was calculated for $\beta = 0.04, 0.20, 0.37, 0.59, 0.79, 0.96, 1.15, 1.35, 1.58, 1.74$, and 1.82. These sequences were chosen for analysis as: (1) the mean β_{opt} value found in Experiment 1 (1.38) is close to the overall β_{opt} of ≈ 1.50 found from Experiments 1 and 2; and (2) the large number of points ($4 \times 2,048$) will give an accurate reflection of the interval distributions associated with each β value. The interval data from the $1/f^\beta$ tone sequences were then correlated (two-tailed) with von Hippel’s (2000) and Vos & Troost’s (1989) data over the interval range 0 to 12 semitones. If the hypothesis made above is correct, we would expect the highest correlations to occur when $\beta = 1.58$.

Figure 6 shows the correlation coefficients plotted against β . For von Hippel’s (2000) data, the highest

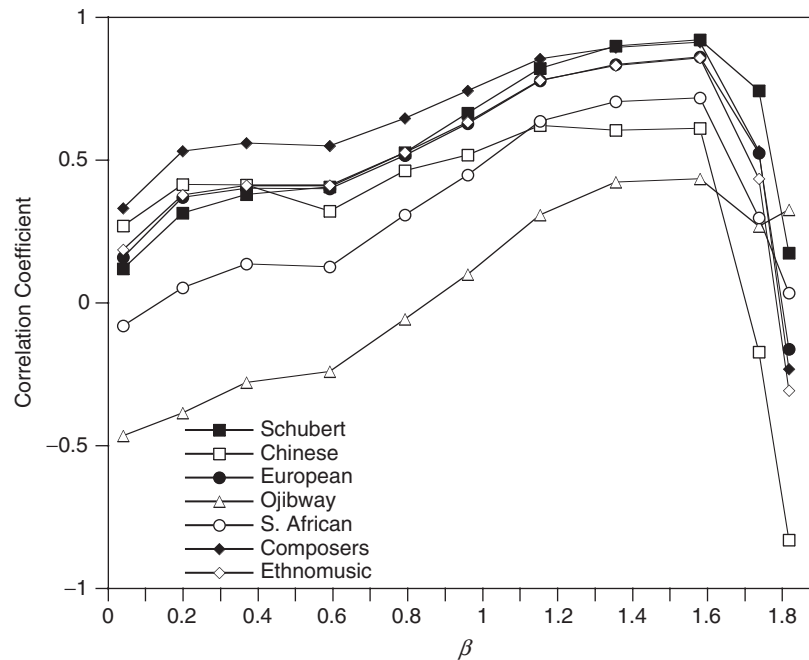


FIGURE 6. Correlation coefficients plotted against β for von Hippel's (2000) data (Schubert, Chinese, European, Ojibway, South African) and Vos and Troost's (1989) data (Composers, Ethnomusic).

correlations were with $\beta = 1.58$ for Schubert, $r(11) = .92$, $p < .001$, European, $r(11) = .86$, $p < .001$, South African, $r(11) = 0.72$, $p = .009$, and Ojibway, $r(11) = 0.43$, NS. However, the highest correlation with Chinese was for $\beta = 1.15$, $r(11) = 0.62$, $p = .023$. For Vos & Troost's (1989) data, the highest correlations were with $\beta = 1.58$ for composers, $r(11) = 0.91$, $p < .001$, and ethnomusic, $r(11) = 0.91$, $p < .001$. As the majority of the highest correlations occurred when $\beta = 1.58$, these findings indicate that the hypothesis made above is substantially correct. That is, the overall β_{opt} of ≈ 1.50 found from Experiments 1 and 2 is valid, and the pitch-interval distribution in a melody determines its perceived complexity and, consequently, the listener's preference. This conclusion also is supported by the plateau for the correlation-coefficient traces in Figure 6, which indicate that the β values of von Hippel's (2000) and Vos & Troost's (1989) melodies mostly lie between 1.35 and 1.58. However, the high correlations for Vos & Troost's (1989) data, and von Hippel's (2000) Schubert and European data, compared to the lower correlations for von Hippel's (2000) South African, Ojibway, and Chinese data suggests that the overall β_{opt} of ≈ 1.50 (that was obtained from UK undergraduates) may only apply to Western/European music, and that β_{opt} may vary across cultures. This notion is supported by: (1) a correlation-coefficient plateau for Chinese melodies from $\beta = 1.15$ to 1.58, indicating a β_{opt} of ≈ 1.35 , and a preference for

melodies with a greater proportion of large pitch intervals compared to Western/European listeners; and (2) higher correlation coefficients when $\beta \geq 1.74$ for Ojibway melodies compared to the others, indicating a preference for melodies with a greater proportion of small pitch intervals compared to Western/European listeners.

General Discussion

Summary of Findings

The results of Experiments 1 and 2 confirmed the three initial hypotheses. With respect to Hypothesis (1), perceived complexity decreased as β increased, indicating that the β value (i.e., spectral power density) of a melody can be used as an objective, quantifiable measure of its perceived complexity. With respect to Hypothesis (2), aesthetic preference (as measured by melodicty) followed an inverted-U function when plotted against objective complexity (as measured by β) and perceived complexity (as measured by the complexity ratings), thereby supporting Schroeder's (1991) association of β with aesthetic preference and the results of previous studies that have found an inverted-U function for music preference when plotted against objective complexity (Crozier, 1974; Vitz, 1966) and perceived complexity (North & Hargreaves, 1995; Orr

& Ohlsson, 2001). Furthermore, the reproduction of the sparse fits found by North & Hargreaves (1995) and Orr & Ohlsson (2001) for the low-complexity sides of the inverted-U functions indicated that: (1) this effect occurs when an independent-subjects design is used, or when the stimuli do not fall over the whole possible range of complexity; (2) $\beta < 2$ for most music stimuli; and (3) ecologically valid melodies only exist over the β range 0.67 to 2.35, with melodies whose β values lie outside this range not being perceived as musical due to their extreme complexity or simplicity. With respect to Hypothesis (3), the overall β_{opt} of ≈ 1.50 found for preferred melodies was between 1 and 2. This finding agrees with the results of Nettheim (1992), Brillinger & Irizarry (1998), and Yadegari (1992), and supports Nettheim's (1992) criticisms of Voss & Clarke's (1975, 1978) methodology. The results of Experiments 1 and 2 also indicated that the perceived complexity of a melody is determined by the distribution characteristics of the pitch intervals between successive notes. Consequently, optimally preferred melodies have a preponderance of small pitch intervals compared to large ones, a characteristic of melodies from a wide variety of cultures (Huron, 2001; von Hippel, 2000; Vos & Troost, 1989). Finally, although it could be argued that the findings reported above, which were obtained using chromatic/ atonal stimuli, may not necessarily apply to music that reflects the cultural norms of participants (e.g., being diatonic, staying in one key, or a related group of keys), this is unlikely to be the case, as the experimental results converged with the findings of studies that used "real" music taken from a variety of cultures (Brillinger & Irizarry, 1998; Nettheim, 1992; North & Hargreaves, 1995; Orr & Ohlsson, 2001; von Hippel, 2000; Vos & Troost, 1989; Yadegari, 1992). Consequently, one would expect "real" music to show the same relationship between β , complexity, melodicty, and pitch-interval distribution that was found for the chromatic/atonal stimuli used in this study.¹

¹Objections to the generalisability of these findings to "real" music might also be raised on the grounds that the stimuli were isochronous. However, as previously noted, Nettheim's (1992) findings indicated that the duration content of a melody makes a minimal contribution to its overall spectrum, with the pitch content being the dominant factor. Furthermore, Nettheim (1992) also noted that the resultant of a white-noise ($\beta = 0$) pitch sequence and a white-noise duration sequence is not necessarily "white-noise music," and reported the result of an informal experiment where he compared two white-noise pitch sequences, the first with random (white noise) pitch durations, the second with uniform pitch durations. He found that the combined white-noise pitch/duration sequence sounded more coherent than the uniform-duration sequence due to the introduction

Stimulus Generation Issues

In the Introduction, we noted the results of two studies that indicated that inverted-U functions are associated with $1/f^\beta$ tone sequences, and that listeners' responses to such sequences reach a maximum at $\beta \approx 2.10$ to 2.40. The first, by Schmuckler & Gilden (1993; Experiment 3), found that the sensitivity of listeners to $1/f^\beta$ tone sequences followed an inverted-U as a function of β , with the peak sensitivity occurring at $\beta \approx 2.40$, and not $\beta = 1$, as would be expected from Voss & Clarke's (1975, 1978) findings. The second, by Patel & Balaban (2000), found that the highest amount of synchronised activity among brain regions in response to $1/f^\beta$ tone sequences occurred when $\beta = 2.10$. Although these findings would appear to conflict with the overall β_{opt} of ≈ 1.50 derived from Experiments 1 and 2, it should be noted that both of these studies used stimuli that were generated by binning $1/f^\beta$ noises into discrete pitch levels. This procedure will result in the aliasing of high frequencies due to quantizing error and, in effect, will add white noise to the spectrum of the original $1/f^\beta$ noise (Eke et al., 2002). With respect to the stimuli, this means that there will be a greater proportion of large pitch intervals in the sequences than expected. With respect to the β values of the stimuli, this means that the spectral slopes of the sequences will be flattened, i.e., the actual β value will be lower than its original value when doing a linear fit to the spectrum on a log/log plot. Therefore, Schmuckler & Gilden's (1993) data, and the inverted-U function they found, should actually be shifted leftward to lower β values to reflect the true nature of the stimuli they used. Consequently, the peak sensitivity will now occur at $\beta < 2.40$. Similarly, for Patel & Balaban's (2000) study, the effect of quantizing error means that the highest amount of synchronised activity actually was generated by sequences where $\beta < 2.10$. The reanalysis of these two studies therefore reduces the conflict of their findings with the overall β_{opt} of ≈ 1.50 found above, and indicates that the spectral slopes of all tone sequences generated from $1/f^\beta$ noises should be checked before use, as in Experiments 1 and 2, to eliminate any effects of quantizing errors that may have occurred in the discretisation process.

of random rhythmic figures. Although no systematic investigation of melodies with combined pitch- and duration-sequences of different β values has been carried out to date, Nettheim's (1992) results indicate this as an area worthy of future exploration. Moreover, the convergence of the findings reported here with those of other studies suggests that, for such stimuli, optimally preferred melodies will result when the combined pitch- and duration-sequences result in a melody where $\beta \approx \beta_{opt} \approx 1.50$.

Neurophysiological Evidence for β_{opt}

For the present sample of participants, the optimal blend of large and small intervals that created the most preferred melodies occurred when $\beta \approx 1.50$. According to Berlyne's (1971) inverted-U theory of aesthetic preference, melodies with this β value should give rise to an optimal level of psychobiological arousal. This suggests that the inverted-U functions for melodicity against complexity and the value of the overall β_{opt} found from Experiments 1 and 2 may have a neurophysiological origin.

One possible neurophysiological mechanism to account for the findings of Experiments 1 and 2 is suggested by research on fractal aesthetics in the visual domain. Spehar, Clifford, Newell, & Taylor (2003) found that aesthetic preference for fractal images follows an inverted-U function against fractal dimension (D), with peak preference occurring around $D = 1.30$. Aks & Sprott (1996) also found that their experimental participants preferred fractal patterns where $D = 1.30$ and, noting that $D = 1.30$ is close to the D value of natural fractal patterns (e.g., waves and clouds), proposed that peoples' preference is set at $D = 1.30$ because of their continual exposure to natural fractal patterns in the environment. A similar proposal was advanced by Knill, Field, & Kersten (1990), who found that the ability to discriminate fractal images followed a quadratic function when plotted against D , with the peak discrimination (i.e., sensitivity) occurring at $D = 2.50$. Similar quadratic functions and D values also were found for the discrimination of fractal images by Gilden, Schmuckler, & Clayton (1993). Knill et al. noted that the D value they found for peak discrimination corresponds to those of natural terrain surfaces, and suggested that the visual system might be optimally tuned to the statistical structure of natural environments. Recently, some physiological evidence to support the optimal-tuning hypotheses of Aks & Sprott and Knill et al. has been found by Yu, Romero, & Lee (2003), who investigated the response of neurons in the primary visual cortex (V1) to moving sine wave grating stimuli with $1/f^\beta$ power spectra. Their results indicated that the maximal response occurred when $\beta = 1$, suggesting that V1 neurons are tuned to optimally process natural temporal signals.

The findings for visual stimuli reported above suggest that the auditory system might also be optimally tuned, but to the statistical structure of natural sound environments, which typically have $1/f^\beta$ power spectra where $\beta \approx 1$ (de Coensel, Botteldooren, & de Muer, 2003). If this is the case, then two hypotheses can be made: (1) the

maximum processing efficiency, sensitivity, and response of the auditory system should occur in response to $1/f^\beta$ signals; and (2) peak performance in response to $1/f^\beta$ signals should occur in the range $\beta \approx 1$ (from de Coensel et al.) to $\beta \approx 1.50$ (from Experiments 1 and 2). In support of Hypothesis (1), studies have shown that $1/f^\beta$ tone sequences reduce EEG complexity (Birbaumer, Lutzenberger, Rau, Braun, & Mayer-Kress, 1994; Jeong, Joung, & Kim, 1998) and increase synchronized activity between brain areas (Patel & Balaban, 2000). In support of Hypothesis (2), Schnupp, Garcia-Lazaro, & Ahmed (2004) have found that neurons in the primary auditory cortex (A1) exhibit tuning to $1/f^\beta$ dynamics and respond maximally when $1 < \beta < 1.50$. Further evidence supporting the optimal-tuning hypothesis advanced above also is provided by a study by Singh & Theunissen (2003) who, for a selection of natural sounds, calculated the changes in their amplitude envelopes with respect to time and frequency (temporal and spectral modulation, respectively). They found that the average temporal modulation power and average spectral modulation power of natural sounds followed $1/f^\beta$ functions when plotted against modulation frequency. That is, the greatest temporal and spectral changes in the amplitude envelopes occurred at the slowest modulations, and these changes decreased in strength with increasing modulation frequency as a function of $1/f^\beta$. Singh & Theunissen suggested that the auditory system has evolved to process natural sounds, and cited neurophysiological evidence indicating a match between the modulation spectra they found and the spectrotemporal receptive fields of auditory neurons (Grace, Amin, Singh, & Theunissen, 2003; Miller, Read, Escabi, & Schreiner, 2002). Singh & Theunissen's results also indicated that: 1) for the average temporal components of the modulation spectra, $\beta = 2.26$ for zebra-finch song, $\beta = 1.60$ for speech, and $\beta = 0.78$ for environmental sounds; and 2) for the average spectral components of the modulation spectra, $\beta = 1.0$ for zebra-finch song, $\beta = 1.52$ for speech, and $\beta = 1.40$ for environmental sounds. The close match between the overall $\beta_{opt} \approx 1.50$ found from Experiments 1 and 2 and the β values for the modulation spectra of speech suggests that preference for melodies where $\beta \approx 1.50$ may be a consequence of the auditory system being optimally tuned to the statistical properties of speech due to continued exposure in the natural environment. If this hypothesis is correct, then it implies that the results described earlier in this paper, which indicated that β_{opt} varies across cultures, are due to the β_{opt} for a culture's music being determined by the characteristics of its language. For example, cultures with languages that resulted in low

β_{opt} values in listeners should have music with high interval-SDs and a greater variability in pitch intervals compared to cultures with languages that resulted in a high β_{opt} values in listeners. Evidence to support this hypothesis has recently been found by Patel, Iversen, & Rosenberg (2006), who showed that the variability of pitch intervals between successive vowels in English and French sentences was reflected in the pitch intervals between successive notes in classical instrumental music by English and French composers, with the pitch-interval variability being greater for English sentences and English music.

The evidence cited above therefore offers some support to the hypothesis that the auditory system is optimally tuned to the statistical structure of natural sound environments, and suggests that the optimal tuning is set to $\beta \approx 1.50$ due to continual exposure to the statistical properties of speech. If this is the case, then the results of Experiments 1 and 2 become explicable. First, listeners will prefer melodies whose statistical structure is close to speech (i.e., the most easily processed melodies), leading to the overall β_{opt} of ≈ 1.50 . Second, for melodies where β is greater than or less than β_{opt} , there will be a drop-off in processing performance that will increase as the difference between β and β_{opt} increases. Consequently, melodies will become harder and harder to process efficiently on either side of β_{opt} and therefore less and less preferred (i.e., less melodic), resulting in the inverted-U functions for melodicity when plotted against β . However, there is also some evidence to suggest that the results of Experiments 1 and 2 may be due to other, more global neurophysiological mechanisms. For example, studies by Watters (1998, 2000), that examined EEG signals taken from

participants under eyes-closed rest, indicated that the EEG signals were a fractal time series with a $1/f^\beta$ spectral density where the mean β across participants was ≈ 1.50 . The convergence of this value with the overall β_{opt} of ≈ 1.50 for melodies suggests that $1/f^\beta$ tone sequences with this β value may be preferred by listeners because of a resonance or entrainment effect created by the stimuli that reproduces in the listener the neuronal activity of the brain when it is in a resting state. Some support for such a resonance effect is given by Patel & Balaban's (2000) finding that the highest amount of synchronised activity among brain regions in response to $1/f^\beta$ tone sequences occurs when $\beta = 2.10$ (actually, $\beta < 2.10$, as described above). In conclusion, although the evidence given above does not allow a definitive biological explanation of the experimental findings reported in this paper, it nonetheless suggests that neurophysiological factors associated with the auditory system, as well as cultural/environmental factors, are at least partly responsible and, consequently, indicates some avenues for future research.

Author Note

Thanks go to two anonymous reviewers for their comments on earlier versions of this paper, Nigel Nettheim for data, and Jens Timmer for FORTRAN code.

Correspondence concerning this article should be addressed to: Michael W. Beauvois, 104 Ratcliffe Road, Loughborough, Leics, LE11 1LH, United Kingdom. E-MAIL: mwbeauvois@ntlworld.com

References

- AKS, D., & SPROTT, J.C. (1996). Quantifying aesthetic preference for chaotic patterns. *Empirical Studies of the Arts*, 14, 1-16.
- ARNHEIM, R. (1985). The other Gustav Theodor Fechner. In S. Koch & D.E. Leary (Eds.), *A century of psychology as science* (pp. 856-865). New York: McGraw-Hill.
- BERLYNE, D.E. (1971). *Aesthetics and psychobiology*. New York: Appleton-Century-Crofts.
- BIRBAUMER, N., LUTZENBERGER, W., RAU, H., BRAUN, C., & MAYER-KRESS, G. (1996). Perception of music and dimensional complexity of brain activity. *International Journal of Bifurcation and Chaos*, 6, 267-278.
- BIRKHOFF, G.D. (1933). *Aesthetic measure*. Cambridge, Mass.: Harvard University Press.
- BRILLINGER, D.R., & IRIZARRY, R.A. (1998). An investigation of the second-and higher-order spectra of music. *Signal Processing*, 65, 161-179.
- CROZIER, J.B. (1974). Verbal and exploratory responses to sound sequences varying in uncertainty level. In D.E. Berlyne (Ed.), *Studies in the new experimental aesthetics: Steps towards an objective psychology of aesthetic appreciation* (pp. 27-90). New York: John Wiley.
- DE COENSEL, B., BOTTELDOOREN, D., & DE MUER, T. (2003). $1/f$ noise in rural and urban soundscapes. *Acta Acustica united with Acustica*, 89, 287-295.
- EKE, A., HERMAN, P., KOCSIS, L., & KOZAK, L.R. (2002). Fractal characterization of complexity in temporal physiological signals. *Physiological Measurement*, 23, R1-R38.

- GARDNER, M. (1978). Mathematical games: White and brown music, fractal curves and one-over-f fluctuations. *Scientific American*, 238, 16-32.
- GILDEN, D.L., SCHMUCKLER, M.A., & CLAYTON, K. (1993). The perception of natural contour. *Psychological Review*, 100, 460-478.
- GRACE, J.A., AMIN, N., SINGH, N.C., & THEUNISSEN, F.E. (2003). Selectivity for conspecific song in the zebra finch auditory forebrain. *Journal of Neurophysiology*, 89, 472-487.
- HENDERSON-SELLERS, B., & COOPER, D. (1993). Has classical music a fractal nature? A reanalysis. *Computers and the Humanities*, 27, 277-284.
- HSU, K.J., & HSU, A.J. (1990). Fractal geometry of music. *Proceedings of the National Academy of Sciences of the USA*, 87, 938-941.
- HSU, K.J., & HSU, A.J. (1991). Self-similarity of the 1/f noise called music. *Proceedings of the National Academy of Sciences of the USA*, 88, 3507-3509.
- HURON, D. (2001). Tone and voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception*, 19, 1-64.
- JEONG, J., JOUNG, M.K., & KIM, S.Y. (1998). Quantification of emotion by nonlinear analysis of the chaotic dynamics of EEGs during perception of 1/f music. *Biological Cybernetics*, 78, 217-225.
- KNILL, D.C., FIELD, D., & KERSTEN, D. (1990). Human discrimination of fractal images. *Journal of the Optical Society of America*, 77, 1113-23.
- MILLER, L.M., ESCABI, M.A., READ, H.L., & SCHREINER, C.E. (2002). Spectrotemporal receptive fields in the lemniscal auditory thalamus and cortex. *Journal of Neurophysiology*, 87, 516-527.
- NETTHEIM, N.N. (1992). On the spectral analysis of melody. *Interface*, 21, 135-148.
- NORTH, A.C., & HARGREAVES, D.J. (1995). Subjective complexity, familiarity, and liking for popular music. *Psychomusicology*, 14, 77-93.
- NORTH, A.C., & HARGREAVES, D.J. (1997). Liking for musical styles. *Musicae Scientiae*, 1, 109-128.
- ORR, M.G., & OHLSSON, S. (2001). The relationship between musical complexity and liking in jazz and bluegrass. *Psychology of Music*, 29, 108-127.
- PATEL, A.D., & BALABAN, E. (2000). Temporal patterns of human cortical activity reflect tone sequence structure. *Nature*, 404, 80-84.
- PATEL, A.D., IVERSEN, J.R., & ROSENBERG, J.C. (2006). Comparing the rhythm and melody of speech and music: The case of British English and French. *Journal of the Acoustical Society of America*, 119, 3034-3047.
- SAUPE, D. (1988). Algorithms for random fractals. In H.O. Peitgen & D. Saupe (Eds.), *The science of fractal images* (pp. 71-113). New York: Springer-Verlag.
- SCHMUCKLER, M.A., & GILDEN, D.L. (1993). Auditory perception of fractal contours. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 641-660.
- SCHNUPP, J.W.H., GARCIA-LAZARO, J.A., & AHMED, B. (2004). Tuning to natural stimulus dynamics in primary auditory cortex. *Society for Neuroscience Abstracts*, 31, 987.6.
- SCHROEDER, M.R. (1991). *Fractals, chaos, power laws: Minutes from an infinite paradise*. New York: W.H Freeman.
- SINGH, N.C., & THEUNISSEN, F. E. (2003). Modulation spectra of natural sounds and ethological theories of auditory processing. *Journal of the Acoustical Society of America*, 114, 3343-3353.
- SPEHAR, B., CLIFFORD, C.W.G., NEWELL, B. & TAYLOR, R.P. (2003). Universal aesthetic of fractals, *Computers and Graphics*, 27, 813-820.
- TIMMER, J., & KÖNIG, M. (1995). On generating power law noise. *Astronomy and Astrophysics*, 300, 707-710.
- VITZ, P. C. (1966). Affect as a function of stimulus variation. *Journal of Experimental Psychology*, 71, 74-79.
- VON HIPPEL, P. (2000). Redefining pitch proximity: Tessitura and mobility as constraints on melodic intervals. *Music Perception*, 17, 315-327.
- VOS, P.G., & TROOST, J.M. (1989). Ascending and descending melodic intervals: Statistical findings and their perceptual relevance. *Music Perception*, 6, 383-396.
- VOSS, R.F., & CLARKE, J. (1975). 1/f noise in music and speech. *Nature*, 258, 317-318.
- VOSS, R.F., & CLARKE, J. (1978). 1/f noise in music: Music from 1/f noise. *Journal of the Acoustical Society of America*, 63, 258-263.
- WATTERS, P.A. (1998). Fractal structure in the electroencephalogram. *Complexity International*, 5. Retrieved April 12, 2006 from <http://journal-ci.csse.monash.edu.au/ci/vol05/watters/watters.html>
- WATTERS, P.A. (2000). Time-invariant long-range correlations in electroencephalogram dynamics. *International Journal of Systems Science*, 31, 819-825.
- YADEGARI, S.D. (1992). Self-similar synthesis on the border between sound and music. Unpublished master's thesis, MIT. Retrieved April 12, 2006 from <http://crca.ucsd.edu/~syadegar/MasterThesis/>
- YU, Y., ROMERO, R.D., & LEE, T.S. (2003). Preference of V1 neurons to natural power spectrum. *Program No. 338.2*. Online Abstract Viewer/Itinerary Planner. Society for Neuroscience: Washington, DC.