

## Perception of Vicentino's 31-tone tuning system

Stephen McAdams<sup>1</sup>, Mikaela Miller<sup>1</sup>, Jonathan Wild<sup>1</sup> and Bruno L. Giordano<sup>2</sup>

<sup>1</sup> Schulich School of Music, McGill University

<sup>2</sup> Institute of Neuroscience and Psychology, University of Glasgow

**Background in music theory.** Nicola Vicentino, a 16<sup>th</sup>-century Italian music theorist and composer, developed a novel 31-tone tuning system to accommodate his adapted theories of the Ancient Greek genera. Vicentino was also interested in the emotional effects that the Ancients claimed music could produce, and he believed that his 31-tone tuning system would be capable of achieving affective power. However, his compositions were never widely performed and fell into obscurity.

**Background in music perception.** Previous studies in tuning system perception, pitch discrimination and musical interval perception indicate that the acoustical differences between nearly identical musical passages in 12-tone equal temperament (12-TET) and 31-tone equal temperament (31-TET) should be discriminable to musically trained listeners. However, as several authors have postulated, there is an effect of context on the perception of musical events. Thus, acoustical differences that should be detectable according to experimentally determined thresholds of tones in isolation may not be predictably perceived in complex musical settings.

**Aims.** This music-theoretical and perceptual study of the microtonal compositions of the 16<sup>th</sup>-century composer and theorist Nicola Vicentino investigates the ability of trained musicians to discriminate between acoustically and musically complex performances of two-chord musical excerpts composed by Vicentino. The extent to which the voice-leading and harmonic features of the music may affect discrimination performance is examined.

**Main contribution.** The music analyzed and used as stimulus materials for the perceptual experiments was taken from Vicentino's (1555/1996) treatise *Ancient Music Adapted to Modern Practice*. Evidence from the two experiments suggests that musically trained listeners can generally detect the differences between 12-TET and 31-TET in an ABX discrimination task. The data were analyzed using generalized linear mixed models. An analysis of the musical correlates suggests that discrimination performance is affected by musical context, and post-hoc analyses demonstrate that excerpts with identical voice-leading contexts elicit similar discrimination performance, whereas excerpts with equal acoustical differences elicit significantly different discrimination responses. The results of a second experiment show that trial structure also affects discrimination performance with end chords having more weight in the comparison.

**Implications.** Music theorists have developed various measures that quantify and compare voice-leading patterns. Many of these measures were developed without regard to perception or the tuning system involved. Vicentino's 31-tone compositions offer the opportunity to relate some of these voice-leading measures to the perception of a novel tuning system.

**Keywords:** Nicola Vicentino, 31-tone equal temperament, 12-tone equal temperament, voice leading, discrimination.

• *Correspondence:* Stephen McAdams, Schulich School of Music, 555 Sherbrooke Street West, Montreal, Quebec, Canada, H3A 1E3; tel. 514-398-4535 ext.094827; e-mail: [smc@music.mcgill.ca](mailto:smc@music.mcgill.ca)

• *Received:* 03 July 2014; *Revised:* 7 October 2015; *Accepted:* 7 October 2015

• *Available online:* 30 October 2015

• doi: 10.4407/jims.2015.10.002

## Introduction

For centuries, tuning systems were central to music theoretical discussions as they laid out the raw materials for composition. However, surviving compositions that explicitly require a specific tuning system are rare. Nicola Vicentino's (1555/1996) treatise *Ancient Music Adapted to Modern Practice* provides an original theory of intervals and tuning, as well as compositions that strive to embody his theory. Vicentino's musical philosophy involves the listener's emotions and reason and argues that perception is an important consideration in the development of theories and in writing music.

Vicentino's compositions offer a unique opportunity to investigate perception in the context of a novel tuning system and elucidate what kinds of musical choices and contexts best highlight the nuances of his system. The present paper reports the results of two perceptual studies that investigated the ability of musically trained listeners to discriminate between short musical excerpts for voices presented in Vicentino's 31-tone tuning system and in the standard 12-tone equal temperament. Modern tuning software has made it possible to have natural voices singing, but to precisely control the tuning of each note, all the while leaving the natural variation of jitter and vibrato if needed. Musical and acoustical correlates were also examined to determine whether musical context had an effect on discrimination performance.

### Vicentino's musical landscape

Renaissance musicologist Edward Lowinsky (1967) acknowledges the innovation found in Vicentino's works, referring to Vicentino as "the greatest revolutionary of the sixteenth century" (p. 133). The novelty of Vicentino's ideas extends beyond his theoretical treatment of the rules of counterpoint and voice-leading, the adaptation of ancient Greek genera to modern composition, and his 31-tone tuning system that can accommodate versions of the Greek diatonic, chromatic, and enharmonic genera.<sup>1</sup> Vicentino's philosophy emphasizes the expressive and emotion-inducing powers of music, and states that sense and reason were equally important in the perception of music (p. 6). The notion that music was able to induce emotions in listeners was shared by many of Vicentino's contemporaries, and dates back at least as far as Ancient Greece (Barker, 1984; Maniates, 1996). Although Vicentino did adapt ancient theoretical ideas in his treatise, his ultimate interest was in resuscitating the affective powers of music as documented by the Greeks, not in reviving the ancient music itself. Throughout his treatise, Vicentino articulates the belief that music has affective powers drawn out through compositional choices (pp. 149-150). Vicentino believed that if a composer of vocal music could capture the essence of the text to be set to music and reproduce it in musical terms, then the resulting composition should have the power to evoke an emotional response in the listener. The proper expression of text, Vicentino claimed, is achieved primarily through the colourings and inflections achieved by employing a broader range than usual of melodic intervals. The realization of these intervals in time as one note moves to the next—the voice-

leading—is of the foremost importance in Vicentino’s compositional theory and more broadly in Renaissance music theory (Bent, 2002; Berger, 1980; McKinney, 2009; Vicentino, 1555/1996, p. 52). Some more of Vicentino’s theoretical concerns, and the resulting analytic perspectives they provide for his compositions, are given in Miller (2011) and Wild (2014) with online sound examples in the latter paper.

The tuning system conceived by Vicentino divides the octave into thirty-one equal parts; we shall call it “thirty-one-tone equal temperament” or 31-TET. This system arises naturally when a quarter-comma meantone, which tempers (narrows) the perfect fifth to obtain a just major third, is taken to its logical extreme (cf. Wild, 2014, for more detail). It results in much finer gradations of pitch than are found in our familiar 12-tone equal temperament.<sup>2</sup> Vicentino notates the pitches of his 31-tone gamut by employing a system of accidentals (Fig. 1): sharps indicate that a note is raised by a chromatic semitone ( $2/5$  of the 31-TET whole tone), flats indicate that a note is lowered by a chromatic semitone ( $2/5$  of the 31-TET whole tone), and a dot above the note indicates that it is raised by  $1/31$  of an octave, equivalent to one fifth of a 31-TET whole tone (hereafter referred to as a fifth-tone, e.g. C vs.  $\dot{C}$ ). Vicentino’s compositions largely employ chords that we recognize today as major and minor chords, some uninflected by dots and others where all notes appear with dots. The chords that consist of dotted notes in 31-TET, then, are inflected sharpwards by one fifth of a tone and thus contain pitches substantially higher than their undotted 31-TET and 12-TET counterparts (see Fig. 1). When a dotted pitch is preceded or followed by an undotted pitch in 31-TET, the melodic intervals between adjacent pitches are also necessarily widened or narrowed by a fifth-tone. Thus,  $C\#$  and  $D\flat$  are different pitches with  $D\flat$  being higher than  $C\#$ ;  $D\flat\flat$  is the same pitch as  $\dot{C}$  (two chromatic semitones, or four fifths of a tone below D—equivalently, one fifth of a tone above C),  $\dot{C}\#$  is equivalent to  $D\flat$ ,  $C\times$  is equivalent to  $D\flat$ , etc.

In his 31-tone system, Vicentino explicitly aimed to introduce a greater variety of intervals into contemporary music and to employ the novelty of intervals as a compositional resource. He thereby hoped to be able to recreate the powerful effects of the music of the Ancient Greeks, whose melodic systems involved an elaborate repertoire of finely graded interval sizes (see Wild, 2014, for an in-depth discussion). To illustrate his theory of affect and the palette of intervals available in his 31-tone tuning system, Vicentino included excerpts from several polyphonic vocal compositions in Book III of his treatise. Examples of 31-tone writing in the treatise appear only in *Soav’e dolc’ardore*, *Dolce mio ben*, *Madonna il poco dolce*, and *Musica prisca caput*. Indeed, these are the only surviving examples of 31-tone composition by Vicentino (with the exception of a few short fragments reproduced in Bottrigari’s treatise *Il melone secondo*). Of these four compositions, only *Musica prisca caput* is complete. The historical record shows that Vicentino found it difficult to train his singers to give precise accounts of his music, and he had to accompany them on a specially invented harpsichord, the *archicembalo*, on which all 31 pitches from his system were spread across two manuals. Vincenzo Galilei, father of the scientist Galileo, writes that he heard Vicentino performing his 31-tone music in many cities in Italy, but that it was not received very well (Galilei, ca. 1586, quoted in Palisca, 1954; Vicentino, 1555/1996). More details can be found in Wild (2014).

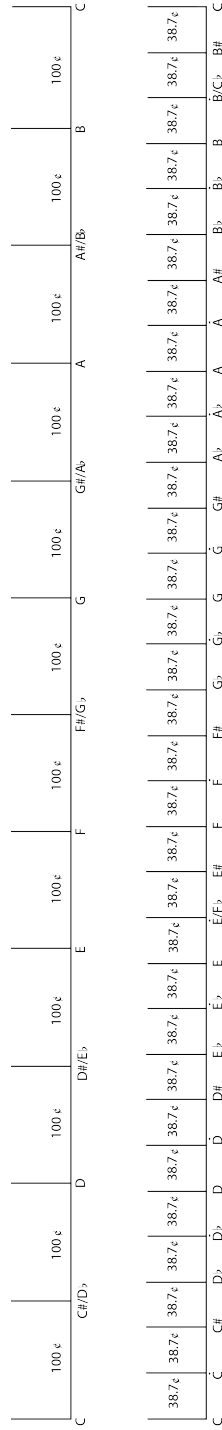


Figure 1. Division of the octave and associated notation. 12-TET (above), and 31-TET (below).

### **Pitch and interval perception**

Tuning systems detail the arrangement and spacing of the members of a gamut of pitches. The spacing between the members of a gamut determines the musical intervals of the system. The discrete pitches are determined once the intervals are mapped to specific points on the pitch continuum. A given tuning system thus implies a set of musical intervals of more or less predetermined magnitudes depending on the culture. Therefore, pitch perception and musical interval perception are both linked to the perception of musical tuning systems.

Pitch discrimination studies investigate the participants' ability to detect whether or not two or more stimuli are identical in terms of frequency. Over the range of frequencies used in music (approximately C0 to B8), even listeners of average ability are able to discriminate more than a thousand different frequencies (Burns, 1999). Pitch discrimination ability is most acute between 500 and 2,000 Hz (approximately B4 to B6) and is very finely tuned; differences of the magnitude of a few cents (1/100 of a semitone) are easily perceived. The minimum discriminable difference in frequency is known as the discrimination threshold. Moore and Moore (2003) presented listeners with successive complex tones with fundamental frequencies between 100 and 400 Hz (approximately G2 to G4, the same range found in Vicentino's vocal music). In each frequency range, listeners had to tell which of two tones had a higher pitch and the difference in fundamental frequency was varied using an adaptive staircase procedure. The pitch discrimination threshold they used was the point at which listeners could just tell the difference 79% of the time. Thresholds fell between 4.3 cents and 20.8 cents depending on fundamental frequency and listener. The difference in absolute pitch height between a dotted note in 31-TET and its 12-TET counterpart always exceeds 22.6 cents and is thus larger than the upper end of the range of thresholds determined by Moore and Moore. Therefore, all dotted notes should be distinguishable from their 12-TET versions based on absolute pitch height alone by normal-hearing listeners in a laboratory setting.

Studies in relative pitch perception have investigated the ability of humans to discriminate, identify and estimate musical intervals, regardless of the absolute frequencies of the stimuli. Moran and Pratt (1926) first determined an average just-noticeable difference (JND) of 18 cents for eleven purely tuned harmonic intervals in an adjustment task. Houtsma (1968) determined JNDs for the 11 melodic intervals of the chromatic scale for one participant, ranging from 13 to 26 cents. Burns and Ward (1978) estimated a mean JND of 37.7 cents for various quarter-tone intervals in a discrimination task. Rakowski (1990) and Rakowski and Miskiewicz (1985) estimated discrimination thresholds between 20 and 45 cents. Burns and Campbell (1994) determined that trained musicians were able to differentially identify intervals separated by 30 cents with an accuracy of about 70%. This finding contradicts Burns and Ward (1978), who concluded that musicians could not reliably use quarter-tone labels separated by 50 cents, but the experimental designs of the two studies differed. In an adjustment task, Burns and Campbell found that their musically trained participants produced standard deviations of 18.2 cents for the chromatic-semitone

intervals and 20.9 cents for quarter-tone intervals. These were not significantly different from each other, leading the authors to conclude that there is evidence that trained musicians can perceive and adjust quarter-tone intervals as accurately as semitone intervals.

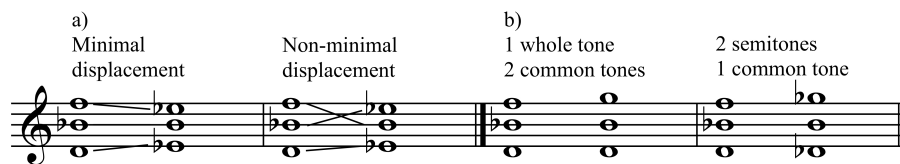
However, Hall and Hess (1984), Rakowski and Miskiewicz (1985), and Vos and van Vianen (1985) all concluded that JNDs are dependent on the interval type; intervals such as the major third, perfect fourth, perfect fifth, octave, and unison have very low discrimination thresholds. The estimated threshold for discrimination of melodic sequences is around 20 cents for an ABX discrimination task (is the X tone the same as the A tone or the B tone?) and 16 cents for an AX discrimination task (is X the same as or different from A?) (Ward and Martin, 1961). The differences in size between these 12-TET melodic intervals and their widened or narrowed 31-TET counterparts are well above these thresholds, and these interval motions are common in Vicentino's music. Therefore, progressions between undotted and dotted chords (or vice versa) should contain discriminable differences in absolute pitch height and relative melodic interval sizes in each of the voices.

Given that we will be exploring the discrimination of chord sequences, two factors need to be considered: the degree to which voice-leading affects the perceived similarity of successive chords and the effects of a musical context on pitch discrimination itself.

### **Voice-leading and the perception of similarity between chords**

A great deal of recent work in music theory has focused on the role of *voice-leading*, which denotes the way in which notes from one chord form connections to the notes in the following chord according to the voice in which the composer has placed them. As noted by Rogers and Callender (2006), various "voice-leading metrics" have been considered as candidate measures for quantifying the distance between chords, with special attention to voice-leading that is *parsimonious*, or smooth, wherein each voice performs the minimal displacement necessary to reach a pitch in the following chord. This is illustrated in Figure 2a. The voice-leading metric that is most commonly used in attempts to quantify chord similarity is the taxicab (or city-block) metric: the linear sum of all the distances traversed by individual voices participating in a succession of two chords. Another approach to chord distance is to count the number of common tones shared by two chords. Rogers and Callender (2006) reported the results of perceptual experiments that investigated listeners' judgments of voice-leading distance between trichords (chords with three pitch classes) and the validity of common-tone retention and total displacement metrics. Their findings showed that the linear "taxicab" metric (in which each semitone movement equates to one unit of distance) was flawed in light of perceived distance, because a movement in a single voice by a whole tone with two common tones retained was rated as being less distant than movement by semitone in two voices while only one common tone was retained (total displacements being 2 semitones in each case, see Fig. 2b). This result suggests that listeners seem to give more weight to common-tone retention than to total displacement in making their judgments. Note that taxicab distance ignores the

contribution of common tones to the perceived similarity of chords. Rogers and Callender further showed that when quarter-tones and 1/8-tones are used, the opposite is true: two voices moving by an 1/8-tone while one common tone is retained was less distant than a single voice moving by one quarter-tone while two common tones were retained (total displacement being 1 quarter tone in each case). Rogers and Callender speculate that this is due to categorical perception: “microtonal tunings may be perceived as alterations of a single pitch rather than as motions from one pitch to another” (p. 1689). In other words, an eighth of a tone is perhaps too small to demand a new interval category, but a quarter tone is not. Their results show that voice-leading metrics do not accurately reflect the perception of voice-leading, and that perceptions of distance can change as interval sizes change.



**Figure 2.** a) Different ways of connecting voices with minimal and non-minimal displacement. b) Chord pairs with similar total displacement (2 semitones) and different common-tone retention (2 tones on the left, 1 tone on the right).

### Effects of musical context on pitch perception

Research has shown that although pitch discrimination of isolated tones is quite good, and likely to be better than the small differences present in Vicentino's 31-TET system, the context in which tones occur can affect discrimination ability. Deutsch (1973), for example, has shown that pitch recognition and pitch discrimination are disrupted when the test tone and the target tone are separated by intervening pitches that are semitones above or below the target pitch or the same as the target pitch. As much of the voice-leading in the excerpts is smooth and involves movement by tones, semitones and fifth-tones, an effect of “pitch interference” might dampen listener’s abilities to properly recognize or discriminate pitches.

Francès (1958) and Rakowski (1990) have both shown that perception can be influenced when stimuli are presented in a musical context. Francès showed that mistunings in the same direction as a strongly expected resolution (i.e., leading tone resolving to the tonic) are not as noticeable to trained listeners, and that mistunings in the opposite direction of a strongly expected resolution are much more noticeable. Rakowski found, however, that trained musicians exhibited decreased variability in their tuning of notes when the note was followed by an expected resolution in an adjustment task. Although the tasks in Rakowski (1990) and Francès (1958) differ, a clear influence of context is shown. As Bigand and Tillmann (2005) suggest, these context effects are so strong that the cognitive system may override sensory processes and fail to accomplish a correct analysis of the situation. They also suggest that the processing of incoming events can be facilitated by expectations; deviant or incoherent events will stand out in juxtaposition to expected events. This suggests that

these context effects will likely affect the perception of the stimuli in this experiment, as the excerpts are taken from real musical passages. The analyses of the possible musical correlates that enhance or detract from discrimination performance will elucidate some of these context effects.

## Notation

The excerpts used in the two experiments each contain a two-chord progression. The progressions were realized such that the intervals of the chords corresponded to those of 31-TET or 12-TET. Each chord was either undotted or dotted (i.e., the undotted chord raised by a fifth-tone). Undotted chords are notated U, and dotted chords are notated D. If the intervals of these chords correspond to those of 12-TET instead of those of 31-TET, they are notated with a prime: U' and D', so U' chords contain normal 12-TET pitches and for D' all pitches are raised a fifth-tone. Without the prime they correspond to 31-TET: U and D. Chords U, D and U' are used in Experiment 1, and all four chord types are used in Experiment 2. In each trial in both experiments, a 31-TET progression is compared to a 12-TET progression. There are three types of comparisons in Experiment 1: DU vs. U'U' (labelled DU), UD vs. U'U' (labelled UD), and UU vs. U'U' (labelled NE for neither chord being dotted). There are four types of comparisons in Experiment 2: DU vs. U'U' (labelled DU), UD vs. U'U' (labelled UD), DU vs. D'D' (labelled DUa), and UD vs. D'D' (labelled UDa).

## Experiment 1

Experiment 1 investigated the extent to which the differences in pitch height and interval size between 12-TET and 31-TET are discriminable in a complex polyphonic musical setting, and examined a set of musical parameters that may be associated with improved discrimination performance.

### Method

**Participants.** Participants were recruited from McGill University and the Montréal area and were compensated for their participation. All participants were highly trained musicians; they were required to be in at least their second year of university-level music studies. Participants also had normal hearing and were regularly practicing musicians. Percussionists and keyboardists were not selected for participation in this study because their extensive practice with a fixed-pitch instrument could, potentially, lead to a lower sensitivity to tuning differences than for variable-pitch instrumentalists and vocalists. There were 16 male participants and 14 female participants, averaging 22 years of age ( $SD=2.9$ ), with 15 years of musical training ( $SD=3.8$ ).

**Stimuli.** The stimuli for the experiment were excerpts taken from recordings of four polyphonic compositions by Nicola Vicentino: *Dolce mio ben*, *Musica prisca caput*, *Soav'e dolc'ardore*, and *Madonna, il poco dolce*. The recordings were made by Jonathan Wild and Peter Schubert at McGill University. Four professional singers (soprano, alto, tenor, and bass) were conducted by Peter Schubert to render a reference recording (see Wild & Schubert, 2008, for recording details). The



compositions were performed by the ensemble without observation of the notated dots; the singers observed only the chromatic accidentals. Later, each singer individually recorded their respective parts while listening to the reference recording. The individual tracks were then retuned using the commercial post-production software Melodyne to produce accurate 12-TET and 31-TET versions of each of the compositions (Wild & Schubert, 2008). Middle C (C4, 262 Hz) was used as the reference frequency to which all other pitches were tuned. Melodyne allows the user to attenuate or amplify vibrato (cyclic variation in pitch at approximately 5-7 Hz) and, independently, any unintended pitch drift within a single note. Given that the singers employed performed in an early-music style—i.e., relatively straight, with little vibrato—only a minimal attenuation of vibrato was incorporated in a few passages of the composition. The results are quite naturalistic. The vocal parts were then remixed into a single track, with moderate reverberation incorporated, to produce recreated performances in precise tuning systems.

The stimuli were 30 excerpts of two-chord progressions from the aforementioned recordings. The two versions of each excerpt (31-TET and 12-TET) were to be compared. The excerpts were chosen so that each chord in the progression had the same number of voices (i.e., no voices entered or dropped out of one of the chords of the excerpt). The excerpts were also triadic, with no embellishing tones. Ten excerpts were chosen to fit into one of the categories: dotted-to-undotted (DU, in which the pitches of the first chord of the excerpt in the 31-TET version are raised by a fifth-tone and the pitches of the second chord are not); undotted-to-dotted (UD, in which the pitches of the second chord of the excerpt in the 31-TET versions are raised by a fifth-tone and the pitches of the first chord are not); or neither (NE, essentially equivalent to UU, in which neither chord of the excerpt in 31-TET involved pitches raised by a fifth-tone). The dotted-to-undotted group (DU) consisted of excerpts in which the first chords of the 12-TET and 31-TET versions of the excerpt differed in terms of tuning, and the second chords were tuned more similarly (DU vs. U'U', where U' represents the 12-TET versions of the chords). The undotted-to-dotted (UD) group consisted of excerpts in which the second chords of the 12-TET and 31-TET versions of the excerpt differed in terms of tuning, and the first chords were tuned more similarly (UD vs. U'U'). The neither group (NE) consisted of excerpts in which neither chord of the 31-TET versions was dotted (UU vs. U'U'), and thus tuning differences were very small.

**Apparatus.** Stimuli were digitally stored on an iMac computer and were played through a Grace Design m904 amplifier to Dynaudio Acoustics loudspeakers in an acoustically treated, soundproof booth (IAC Acoustics, model 120-act3). The experimental interface was created using Adesign software by Pierresoft.com. It was displayed on the iMac computer monitor and responses were entered using the computer mouse.

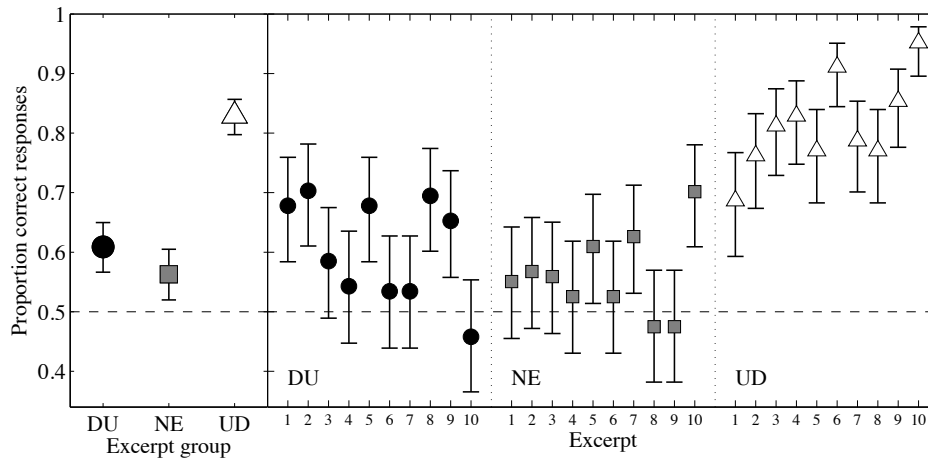
**Procedure.** The musical stimuli were presented in stereo over Dynaudio BM6a loudspeakers at 65 db SPL on average as measured with a Bruel & Kjaer 2205 sound level meter (A-weighting) positioned where the head of the listener would be. Participants were centred between the loudspeakers at a distance of about one meter from them and were seated in front of a computer monitor.

An ABX discrimination task was employed with the stimuli mentioned above. There were 30 pairs of excerpts to be discriminated, with 10 excerpts in each of the DU, UD and NE groups. Each pair consisted of a single excerpt performed in two tunings: once in 12-TET, and once in 31-TET. In each of the 30 trials, the participant was presented with three passages, separated in all cases by a silence of 500 msec. Version A and version B were heard first, and together represented one of the 30 excerpt pairs (one was the 31-TET version of a given excerpt, the other was the 12-TET version). Version X was heard last, and was identical to either version A or version B. The participant was asked to identify whether version X corresponded to version A or version B. A prompt on the computer monitor informed the participant which version they were currently hearing (A, B, or X), and the versions were always played in that order. Once all three versions had been presented, another prompt appeared on the monitor asking the participant to decide whether X was the same as A or B. Participants entered their responses by clicking with the mouse either 'A' or 'B' on the monitor to indicate their choice. Participants could not enter a response until all versions of the excerpt had finished playing. Each trial was played only once, and the participants could not repeat or revisit any of the trials.

A modified block-randomized design was used. In a single block of 30 trials, each of the 30 excerpt pairs occurred once, randomly selected from the four possible experimental orders comparing 12-TET and 31-TET stimuli (12/31/12, 12/31/31, 31/12/12, and 31/12/31), without replacement. The four possible experimental orders were thus randomly distributed across the four blocks, each excerpt occurring only once per block, for a total of 120 trials. It took approximately 10 minutes to complete a block of 30 excerpts.

## **Results**

The goal of this study was to determine if trained musicians could discriminate between musical excerpts in 12-TET and 31-TET when steps of a fifth-tone were involved. Several voice-leading and harmonic features of the excerpts were analysed to determine if they could be correlated with discrimination performance. Figure 3 (left panel) displays the proportion of correct responses by excerpt group with 95% confidence intervals. The horizontal dashed line shows the chance-performance level of 50% correct. If the confidence interval includes this line, we consider that listeners do not reliably detect the difference between 12-TET and 31-TET versions.



**Figure 3.** Proportion of correct responses in Experiment 1 by excerpt group on the left and by individual excerpt on the right. Error bars show the 95% confidence interval about the mean, as derived from the generalized linear mixed model.

Generalized linear mixed models (GLMM) were chosen to analyse the data for two reasons. The first was to accommodate the repeated-measures design, i.e., an experiment in which a given listener makes responses on several different stimuli. This statistical model was chosen over the traditional repeated-measures analysis of variance because it more accurately accounts for the variability between listeners in a repeated-measures design. The second was that our data were not normally distributed due to their being based on a binary outcome (yes/no response). In simple words, GLMMs can be conceived as a regression method that models the effects of an independent variable on the responses of the experimental participants with two types of parameters: the fixed effects, i.e., a measure of the average effect across participants, and the random effects, i.e., a measure of the variability of the participant-specific effects. Differently than in traditional repeated-measures ANOVA, fitting GLMMs requires selecting the effects included in the model (e.g., which fixed and which random effects; West, Welch and Galecki, 2007, pp. 39–41). The models presented here were the result of a forward-selection procedure, i.e., effects were added one by one until the inclusion of additional effects did not result in a significant improvement of the model fit. The final model presented here included a random intercept, modelling the variation across participants in the across-trials average performance, and a random effect of excerpt group, modelling the variability across participants in the effect of excerpt group on performance. Excerpt group (DU, UD, or NE) and individual excerpts nested within excerpt groups were treated as fixed effects in the model, i.e., the main variables of interest that are manipulated experimentally. Within this model, excerpt group (DU, NE, or UD) had a significant effect on performance in the task of discriminating 21-TET and 31-TET systems,  $F(2, 58)=66.7$ ,  $p<0.0001$ , indicating that the three groups had different levels of discrimination performance. The effect of excerpt within group was also significant,  $F(27, 783)=3.4$ ,  $p<0.0001$ , indicating that there was at least one significant difference

in discrimination performance for two excerpts in the same group. Post-hoc contrasts revealed that correct response rates in the UD condition were significantly higher than those of both the DU group,  $F(1, 58)=88.8$ ,  $p<0.0001$ , and the NE group,  $F(1, 58)=122.3$ ,  $p<0.0001$ . Although, on average, the two tuning systems were better discriminated for DU stimuli than for NE stimuli, this difference was only marginally significant,  $F(1, 58)=3.23$ ,  $p=0.08$ , suggesting that the difference in discriminability for DU and NE excerpts was small. Overall, listeners were able to reliably detect the differences in tuning systems even when steps of a fifth-tone were not involved. All three groups taken as wholes elicited above-chance performance as shown in Figure 3 (note that the error bars on the left panel of the graph do not cross the 50% dashed line).

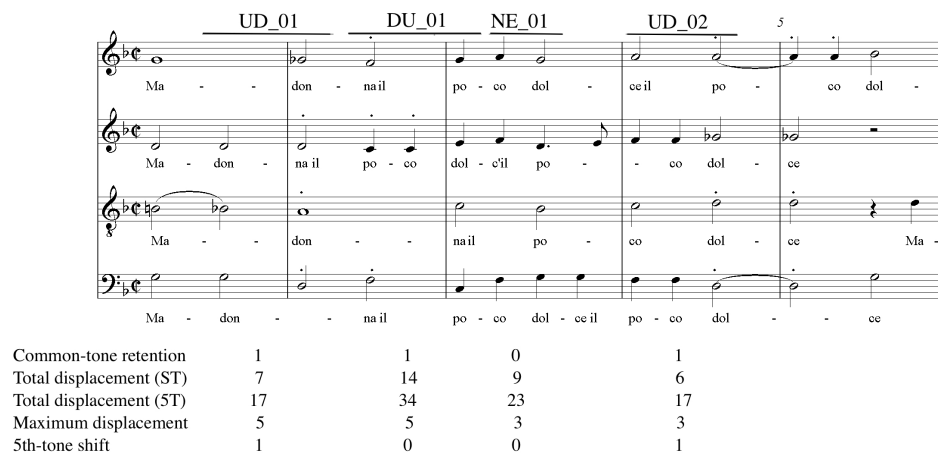
Confidence intervals (95%) were also produced from the GLMM for individual stimuli to investigate which excerpts elicited above-chance performance (Fig. 3, right panel). All of the excerpts involving steps of a fifth-tone (such as the dotted notes found in the DU and UD groups) elicited above-chance performance except DU<sub>4</sub>, DU<sub>6</sub>, DU<sub>7</sub>, and DU<sub>10</sub>. The only NE excerpts that were significantly above chance were NE<sub>5</sub>, NE<sub>7</sub>, and NE<sub>10</sub>.

Using information gathered from questionnaires, the discrimination abilities of the participants were assessed according to their primary instrument type (voice, string, or wind), and whether or not they reported having absolute pitch. It is possible that vocalists might be more sensitive to the tuning in vocal performance and that possessors of absolute pitch might perform differently than non-possessors. We tested these hypotheses by adding either a main-instrument or an absolute-pitch variable to the main non-linear mixed model, which included the fixed effects of excerpt group and of excerpt nested within excerpt group, as well as the random effects of excerpt group and the random intercept. Neither instrument type nor absolute pitch significantly affected performance,  $F(2,783) < 1$  and  $F(1,783) < 1$ , respectively. This finding contradicts Miyazaki (1993), who determined that possessors of absolute pitch performed slightly worse than non-possessors on roaming relative pitch identification tasks, presumably because they rely on absolute pitch relationships for fundamental frequencies that could not be assimilated into one of the 12 chromatic pitch categories.

**Musical parameters.** A group of variables relating to the voice-leading, harmonic, and acoustical properties of the excerpts were tested to investigate which factors were associated with the ability to discriminate 31-TET from 12-TET versions. All of the values for each parameter were standardized by converting to z-scores. This was done so that parameter estimates could be compared in terms of the magnitude of effect size. This GLMM for this portion of the analysis included random effects (slope only) for each listener, and fixed effects for the voice-leading, harmonic, and acoustical parameters. Statistics for predictors with a significant effect on discrimination performance are reported based on the model resulting from a forward selection of musical predictors (again, adding predictors one at a time until new additions don't increase the fit to the data).

*Voice-leading distance.* The excerpts were analysed using various published voice-leading distance metrics. The parameters examined were the sum of the displacements

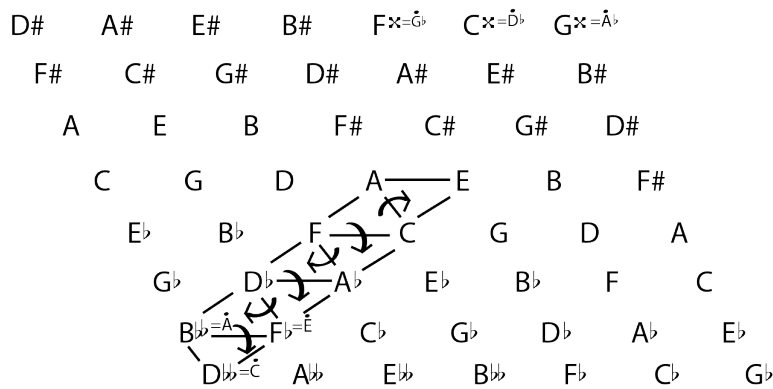
of all the voices in semitones in the 12-TET version and in fifth-tones in the 31-TET version (total displacement, identified as ST and 5T, respectively, in Fig. 4), the maximum displacement of any voice in semitones in the 12-TET versions, the number of pitch classes in common between the two chords when both chords are modified to their 12-TET versions (common-tone retention), the interaction between common tones and total displacement, and whether or not at least one voice in the excerpt realized the step of a fifth-tone, e.g., D to  $\dot{D}$  in the soprano in UD<sub>1</sub> in Fig. 4 (5th-tone shift). Spearman correlations were used to determine the extent to which the rank order of voice-leading distances of one type corresponded to those of another type. This analysis revealed that 12-TET total displacement was highly correlated with maximum displacement,  $r(28)=0.76$ ,  $p<0.0001$ , and with 31-TET total displacement,  $r(28)=0.66$ ,  $p<0.0001$ , indicating that they capture similar things. 12-TET total displacement was thus retained and 31-TET total displacement and maximum displacement were dropped from further modelling. The only voice-leading parameters with a significant effect on performance were total displacement,  $F(1, 864)=20.7$ ,  $p<0.0001$ , and the interaction between total displacement and common-tone retention,  $F(1, 864)=14.8$ ,  $p=0.0001$ . These results suggest that the effect of total displacement on discrimination performance depends on the presence of common tones. This result, in contradistinction to Rogers and Callender's (2006) finding that listeners give more weight to common-tone retention than to total displacement, suggests that an increased degree of pitch movement between chords results in increased discrimination when common pitches are repeated between 12-TET and 31-TET tunings. These results are explored in more detail in Experiment 2.



**Figure 4.** Examples of DU, UD and NE chord pairs from Vicentino's *Madonna, il poco dolce* with corresponding values of voice-leading metrics for number of common tones, total displacement (in semitones for 12-TET version), total displacement (in 5th tones), maximum displacement, and presence of a fifth-tone step in any voice (see text for descriptions of these metrics).

*Harmonic distance.* In Neo-Riemannian theory, harmonies are related directly to each other without reference to a tonic (cf. Cohn, 1998). The parameter tested was the

difference in the harmonic distances between the two pairs of chords of the excerpt, as measured by the minimum number of edge flips to get from one chord to another on the 12-TET Tonnetz and 31-TET Tonnetz. Tonnetze are conceptual diagrams that map pitches in tonal space; successive lines of perfect fifths, major thirds, and minor thirds interconnect to form lattices of triadic relationships (see Fig. 5). In Neo-Riemannian theory, an edge flip signifies the movement from one harmony to an adjacent harmony that contains two common tones. For example, there is one edge flip from F major (F-A-C triangle) to A minor (A-C-E triangle), but there are five edge flips from F major to  $\hat{A}$  minor ( $\hat{A}-\hat{C}-\hat{E}$ , equivalent to  $B\flat$  minor,  $B\flat-D\flat-F\flat$ ). These distances are indicated by the arrows between triads in Figure 5. Increased difference between 12-TET and 31-TET versions of the excerpt in terms of the harmonic distance between the two chords of each version was accompanied by an increased probability of correct discrimination,  $F(1, 864)=9.4$ ,  $p=0.002$ . Thus alterations of closely related chords were more obvious than alterations of distantly related chords.



**Figure 5.** Edge flips for harmonic motion from an F major triad to an A minor triad (harmonic distance of 1, upward to the right) and to an  $\hat{A}$  minor triad (equivalent to  $B\flat$  minor, harmonic distance of 5, downward to the left), illustrated on the 31-TET *Tonnetz*.

*Acoustical parameters.* One parameter examined was the duration (in seconds) of each excerpt. Two other parameters were the sum of the difference in cents between each of the congruent voices (i.e., the difference between the soprano voice, alto voice, tenor voice, or bass voice) in the 12-TET and 31-TET versions. The sum of the difference in cents is a measure of the differences in absolute pitch height of the notes comprising the chord in the 12-TET or 31-TET versions. One parameter computed this sum in the first chord only and the other in the second chord only. The duration of the excerpt had a significant negative effect on performance,  $F(1, 864)=5.1$ ,  $p=0.02$ , with longer excerpts eliciting lower performance (the longest excerpt was 6.0 s; the shortest was 2.6 s). The sum of the difference between the second chords of the excerpt had a highly significant effect,  $F(1, 864)= 56.4$ ,  $p<0.0001$ , but was not significant for the first chords,  $F(1, 863) < 1$ .

Because the values of the voice-leading, harmonic, and acoustical parameters were standardized, the parameter estimates obtained from the analysis models can be interpreted as relative effect sizes. The difference in cents between the second chords of the excerpts had the largest effect on performance, followed by total displacement, the absolute value of the difference in harmonic distance on the Tonnetze, and the duration in seconds.

**Post-hoc analyses.** The excerpts were analysed to determine whether any of them shared the same voice-leading pattern, consisted of the same harmonic chords, or reversed the voice-leading pattern and harmonic pattern of any other excerpt (i.e., one excerpt retraces the voice-leading and harmony of another, thus reversing the directionality of all the intervals and the harmonic progression). The goal of the post-hoc analyses was to determine whether the excerpts with similar voice-leading patterns elicited similar discrimination responses. Excerpts that share the same chords but have different voice-leading patterns would have the same difference in cents between the two versions of the excerpts. Those excerpts that share voice-leading patterns at different pitch levels represent cases in which the voice-leading is held constant, whereas the difference in cents between the two versions of the excerpts varies. And lastly, those excerpts that reverse voice-leading patterns at the same pitch level might show an effect of directionality; if the total difference is constant and the voice-leading patterns are reversed, then the direction of the intervals may be playing a role.

This set of excerpts only offered a few instances of the above-mentioned cases. The excerpts that share the same chords are  $DU_7$  and  $DU_9$ , on the one hand, and  $UD_5$  and  $UD_6$ , on the other hand. The discrimination performance was marginally significant for  $DU_7$  vs.  $DU_9$ ,  $F(1,783)=3.42$ ,  $p=0.065$ , and was significantly different for  $UD_5$  vs.  $UD_6$ ,  $F(1,783)=8.41$ ,  $p=0.0038$ , suggesting that voice-leading patterns may overpower the differences in cents between the versions of the excerpts. The only excerpts to reverse the voice-leading patterns at the same pitch level are  $DU_2$  and  $UD_4$ . A post-hoc comparison<sup>3</sup> revealed that these two excerpts were significantly different in terms of correct-response rates,  $F(1,841)=5.12$ ,  $p=0.024$ , suggesting that directionality may also affect discrimination in a four-voice texture when total difference in cents and voice-leading distance are held constant. This evidence supports the conclusion of Rogers and Callender (2006) that voice-leading is not perceptually linear and symmetric, as is assumed by most metrics for quantifying voice-leading distance. Excerpts  $DU_2$  and  $UD_2$  exhibit the same voice-leading pattern at different pitch-levels, and they are not significantly different in terms of discrimination performance,  $F(1,841)=1.04$ ,  $p=0.31$ .  $DU_2$  is the only DU excerpt to share voice-leading patterns with an excerpt in the highly discriminable UD group. These analyses provide further evidence that the specific voice-leading patterns and musical context may affect discrimination ability.

## Discussion

The results of Experiment 1 confirm the hypothesis that musically trained musicians can reliably tell the difference between 12-TET and 31-TET, in some cases even when the step of a fifth-tone is not involved (three of the ten NE excerpts). A

significant asymmetry was observed between the DU and UD groups; performance was significantly better when participants were discriminating between versions of UD excerpts than when they were discriminating between versions of DU excerpts. The tuning differences between the 12-TET and 31-TET versions are much more pronounced within these groups than within the NE group; the differences in tuning are also similar in the UD and DU groups. Surprisingly, however, mean discrimination performance for the DU group of excerpts was not significantly different from the mean performance for NE excerpts, even though the NE excerpts exhibited smaller differences in tuning between the 12-TET and 31-TET versions.

Of the musical parameters investigated, it was found that total displacement, the interaction between total displacement and common-tone retention, the absolute difference between distances on the 12-TET and 31-TET Tonnetze, the difference in cents between the second chords of the excerpts, and duration (in seconds) of the excerpt had a significant effect on discrimination performance. It was originally thought that the total displacement would have a negative effect on the discrimination performance because a change in interval magnitude of a fifth-tone might be much more noticeable when the interval is small; a fifth-tone is a larger percentage of the entire frequency ratio for a smaller interval than for a larger one. Progressions with large displacements would most likely have several voices moving by leap and might be harder to discriminate. The opposite effect was observed, however. This may be due to the definition of ‘voice’ used in calculating the total displacement. A ‘voice’ was determined by the score; in other words, the voices were the written soprano, alto, tenor, and bass parts. Those excerpts with large total displacements tended to be excerpts with higher common-tone retention, but with large leaps that occurred as the result of a voice exchange. For instance, if tenor and alto both leap a perfect fourth in opposite directions to exchange pitches and avoid parallel perfect intervals (Fig. 6), the chord itself has not changed but the voices have switched relative positions. When this happens in 31-TET and one of the two chords in the excerpts is dotted, all of the pitches are raised by a fifth-tone. This would then be compared to the static 12-TET version in which the pitches do not move. Harmonically static motion in 12-TET compared to voices that move by discriminable amounts in 31-TET elicited higher performance, but tended to have high total displacements as a result of the voice exchanges. The significance of the interaction term provides further evidence to support this hypothesis: when total displacement increases, there is an additional increase in performance as the number of common tones increases. When total displacement was measured by proximity (a ‘voice’ is determined by proximity in pitch, not by written voice-parts), the conclusion did not change.



**Figure 6.** Voice-leading reduction of UD<sub>10</sub>: voice-exchange. (Note that all of the pitches in the second chord are dotted. The dots having been removed to use two staves.)



The absolute difference in harmonic distances as measured by the number of edge-flips on the 12-TET and 31-TET Tonnetze had a notable effect. Harmonic stasis (for example, the repetition of a D-major triad) has a harmonic distance of 0. In a 31-TET environment, a progression from D major to  $\dot{D}$  major has a harmonic distance of 6 ( $\dot{D}$  major, in the guise of  $E\flat\flat$  major, is 6 edge-flips distant from D major). In general, progressions that would be harmonically distant in 12-TET remain harmonically distant in 31-TET. Progressions between undotted and dotted chords (or vice versa) that would be harmonically close if the dots were ignored (i.e., performing the music in a 12-tone environment) are much more distant on the 31-TET Tonnetz. For instance, C major to A minor has a harmonic distance of 1 on the 12-TET Tonnetz, because the triads C major and A minor share an edge; a 31-TET version from C major to  $\dot{A}$  minor has a harmonic distance of 7, due to the fact that we must traverse a path with 7 steps to go from C major to  $B\flat\flat$  minor (for example: C, Cm,  $E\flat$ ,  $E\flat m$ ,  $G\flat$ ,  $G\flat m$ ,  $B\flat\flat$ ,  $B\flat\flat m$ ). Each step involves two triads that share an edge, and thus have two pitch classes in common. A remote 12-TET relationship such as C major to  $F\sharp$  major has a harmonic distance of 4; a 31-TET version from C-major to  $\dot{F}\sharp$ -major is not substantially more remote, exhibiting a harmonic distance of 5.

The difference in absolute distance is thus greater for chords that are harmonically closer in 12-TET, but one of the two chords is dotted in the 31-TET version. The alteration of raising or lowering one of two harmonically close chords by a fifth-tone, then, increased discrimination performance. The 31-TET alteration of what would have otherwise been a harmonically close chord progression is thus more noticeable than alterations of distantly related chords. Rogers and Callender (2006) and Krumhansl (1998) have shown that harmonic distances of 1 on the 12-TET Tonnetz are rated as being perceptually close. The findings of Experiment 1 suggest that tuning alterations to otherwise perceptually close chords are more noticeable.

The duration of the excerpts was found to have a negative effect on discrimination ability. This is most likely due to the limits of short-term memory. Performance in discrimination tasks is highest when the participant is able to rely on long-term and echoic memory cues (Burns, 1999). As the duration of the excerpt increases, it becomes more difficult for the listener to retain an accurate impression of the stimuli in echoic memory. Shorter excerpts make for closer comparisons because the trace impressions of the A, B, and X versions are relatively fresh in memory.

The complexity of the musical stimuli used in this experiment makes it difficult to isolate effects. The excerpts were not normalized for loudness or balance between the voices; they were taken from expressive performances. There may also be an effect of the lyrics; the singers sang the appropriate lyrics for their parts, and so consonant and vowel sounds vary greatly between the excerpts. The engineering of the excerpts may also have had an effect despite measures being taken to reduce differences between the 12-TET and 31-TET versions of the recordings.

The asymmetry of performance for the DU and UD groups, in conjunction with the highly significant effect and large effect size of the difference in cents between the second chords of the excerpts, led to a second experiment that aimed to tease out the extent to which the stimulus presentation affected discrimination performance. If the

participants rely more heavily on a comparison across the versions of the second chords of the excerpt and ignore, to some extent, the first chords of the excerpt, then the nature of the discrimination task may play a role. In DU excerpts, the differences in the absolute pitch height between the 12-TET and 31-TET versions are greater for the first chord than for the second; the opposite is true of the UD excerpts. Experiment 2 investigated the extent to which listeners relied on the absolute pitch differences of the final chords in the discrimination task by retuning the 12-TET versions of each excerpt to alter the pitch height of the final chords.

## Experiment 2

The goal of the second experiment was to determine the extent to which the nature of the trial structure affected discrimination performance. In Experiment 1, we determined that participants were significantly less able to discriminate between the two tuning systems with DU excerpts than with UD excerpts. It was thought that this asymmetry may be attributable to the nature of the trial structure. Because the stimuli are between 2 and 6 seconds in duration, the final chord of each excerpt might have the strongest impression that participants retain in their memory. Thus, when the endings of the excerpts in each trial do *not* match in terms of pitch height (i.e., in the UD vs. U'U' case), it is easier for participants to discriminate between the tuning systems. When the endings 'rhyme', as in the DU vs. U'U', it is more difficult to discriminate between the excerpts because the significant points of comparison—the endings—are extremely similar in terms of pitch height. We therefore control for this in the current experiment by creating a complementary set of stimuli in which the 12-TET versions are tuned up by a fifth-tone. This effectively moves the pitches of the 12-TET scale closer to the corresponding dotted notes of the 31-TET scale, thus alternating the ending pitch height schemes to DU vs. D'D' and UD vs. D'D' (D' represents the 12-TET U' chord raised by a fifth-tone).

## Method

**Participants.** A sample of participants similar to that of Experiment 1 was recruited for Experiment 2 from McGill University and the Montréal area. None of these participants had taken part in Experiment 1. They were compensated for their time. All participants were highly musically trained; they were required to be in at least their second year of university-level music studies. They had to have normal hearing and be regularly practicing musicians. As in Experiment 1, percussionists and keyboardists were excluded. There were 7 male and 8 female participants, averaging 23 years of age ( $SD=5.3$ ) with 16 years of musical training ( $SD=6.1$ ).

**Apparatus.** The apparatus for Experiment 2 was identical to that of Experiment 1.

**Stimuli.** The stimuli for this experiment included 20 of the 30 excerpts used in the Experiment 1, retaining only the UD and DU excerpts. In Experiment 1, participants were asked to compare UD with U'U' and DU with U'U', where U' represents the 12-TET version, which is closer in terms of absolute pitch height to the undotted notes of 31-TET. Experiment 2 also included UD vs. D'D' and DU vs. D'D', where D' represents a U' chord raised by a fifth-tone, which is closer in terms of absolute pitch

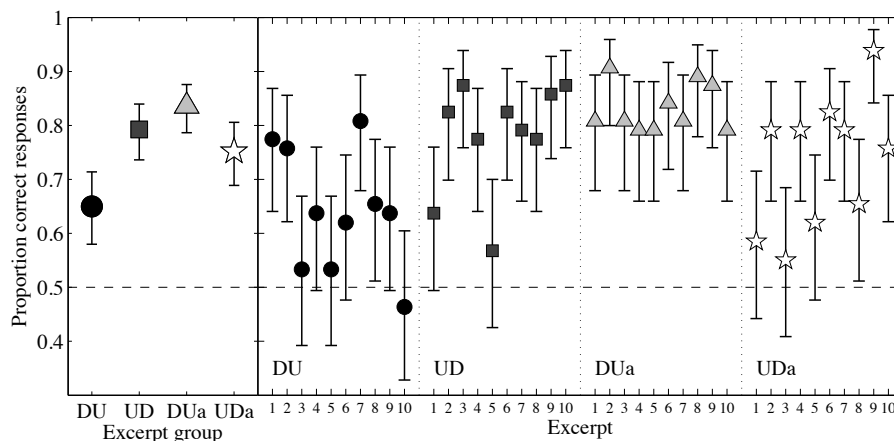
height to the dotted notes of 31-TET. The purpose of producing a D'D' version was to alternate the 'rhyme' schemes of the DU and UD excerpts. Thus there were 40 comparisons to make: twenty 31-TET vs. 12-TET (10 UD vs. U'U' and 10 DU vs. U'U'), and twenty 31-TET vs. 12-TET raised a fifth-tone (10 UD vs. D'D' and 10 DU vs. D'D'). These latter conditions will be identified by UDa and DUa, respectively. The U'U' and D'D' progressions were not directly compared.

**Procedure.** The procedure for Experiment 2 was nearly identical to the one used in Experiment 1. Participants performed the ABX discrimination task with four blocks of 40 excerpts. There were 10 UD excerpts (UD. vs. U'U'), 10 DU excerpts (DU vs. U'U'), 10 DUa excerpts (DU vs. D'D'), and 10 UDa excerpts (UD vs. D'D'). The UD and UDa groups were identical in terms of musical content, as were the DU and DUa groups; they only differed in terms of which 12-TET version (U'U' or D'D') was being compared to the 31-TET version.

Again, a modified block-randomized design was used. In a single block of 40 trials, each of the 40 excerpt pairs occurred once, randomly selected from the four possible experimental orders of 31-TET and 12-TET excerpts (12/31/12, 12/31/31, 31/12a/12a, and 31/12a/31), without replacement. The four possible experimental orders were thus randomly distributed across the four blocks, each excerpt occurring only once per block, for a total of 160 trials. Each block took approximately 15 minutes to complete.

## Results

In Experiment 2, we were interested in determining whether there was an effect of 'rhyming' endings in the presentation of the stimuli. Figure 7 shows the proportion of correct responses for each of the groups of excerpts with 95% confidence intervals. All four groups elicited performance that was significantly above chance, as shown in Figure 7 (left panel).



**Figure 7.** Proportion of correct responses in Experiment 2 by excerpt groups on the left and by individual excerpts on the right. Error bars show the 95% confidence interval about the mean. The horizontal line shows the chance-performance level of 50% correct.

The data were again analysed using a generalized linear mixed model to determine whether there were differences in discrimination performance among excerpt groups or among excerpts within each group, in which excerpt group (DU, UD, DUa, and UDa) and individual excerpts nested within group were treated as fixed effects, and participant was treated as a random effect (intercept only). The model revealed that excerpt group and individual excerpt nested within excerpt group both had significant effects on the odds of correct discrimination,  $F(3, 546)=18.7$ ,  $p<0.0001$  for excerpt group,  $F(3, 546)=2.8$ ,  $p<0.0001$  for effect of excerpt within group. These tests indicate that there were indeed differences in discrimination ability for the different excerpts and groups. The DU excerpt group was significantly lower than all other groups: DUa,  $F(1, 546)=49.3$ ,  $p<0.0001$ ; UD,  $F(1, 546)=27.3$ ,  $p<0.0001$ ; and UDa,  $F(1, 546)=12.7$ ,  $p=0.0004$ . Contrast parameterization revealed that discrimination performance for the UD group was marginally lower than that of the DUa group,  $F(1, 546)=3.5$ ,  $p=0.06$ , but was not significantly different from the UDa group,  $F(1, 546)=2.5$ ,  $p=0.12$ . Performance for the DUa group and the UDa differed significantly, with DUa excerpts being better discriminated than UDa excerpts,  $F(1, 546)=11.6$ ,  $p=0.0007$ . When the DU and UDa (ending 'rhyme') data were contrasted with the UD and DUa (no ending 'rhyme'), a highly significant difference was observed, with rhyming impeding discrimination compared to no rhyming,  $F(1, 546)=36.3$ ,  $p<0.0001$ . Finally, DU and UD contrasted with DUa and UDa revealed a significant difference,  $F(1, 546)=13.2$ ,  $p=0.0002$ , indicating that conditions with a 5th-tone ('a') transposition were better discriminated than those without transposition.

Figure 7 (right panel) displays the proportion of correct responses for each excerpt with 95% confidence intervals. The individual DU excerpts that elicited above-chance performance in Experiment 2 differed slightly from those found in Experiment 1. Of the DU excerpts, DU<sub>1</sub>, DU<sub>2</sub>, DU<sub>4</sub>, DU<sub>7</sub>, DU<sub>8</sub>, and DU<sub>9</sub> elicited above-chance performance. Of the UD excerpts, only UD<sub>5</sub> did not elicit above-chance performance. All of the DUa excerpts elicited above-chance performance, but performance for UDa<sub>1</sub>, UDa<sub>3</sub>, and UDa<sub>5</sub> was not significantly above chance.

The results of Experiment 2 confirm the findings of Experiment 1; overall, UD and DU groups both elicit above-chance performance. In addition, global performance for the UDa excerpts and the DUa excerpts was also significantly different from chance.

## Discussion

The results of Experiment 2 provide evidence to support our hypothesis that there is an ending 'rhyme' effect. Discrimination ability significantly improved for the DU excerpts when the D'D' version was compared to the 31-TET versions. This eliminated the ending rhyme so that the final chord of each version of the excerpt was dissimilar in terms of pitch height. Conversely, the UD performance did not significantly change when a D'D' progression was compared to 31-TET progressions.

The results of Experiment 2 show that listeners can more easily discriminate between the two tuning systems when the pitch heights of the final chords are dissimilar. This effect may be due to the limitations of short-term memory: the excerpts are several seconds in duration with a brief pause between versions, and the last thing heard

leaves the strongest memory trace as it is freshest in the listener's mind (Murdock, 1962). This facilitates the comparison process because the listener is able to recall pitch height more accurately when it is still in recent short-term memory. The effect is not entirely symmetrical; if it were, the size of the effect observed between the DU and DUa groups would be as large as between the UD and UDa groups. Because the improvement in the DUa group was not as equally matched by a drop in performance in the UDa group, it cannot be assumed that the ending effect is symmetrical. Although the absence of an ending rhyme may improve discrimination performance, it is not the only factor. The UDa group has significantly higher odds of correct discrimination than the DU group, even though both groups exhibit the ending rhyme. This result may suggest that there is still some effect of progression type (DU vs. UD), although it is clearly influenced by the comparison context in this experiment (DU/UD vs. DUa/UDa). Although our stimulus sample is too small to evaluate the role of harmonic function, one might note anecdotally that Vicentino's DU successions more often appear as inflections of a root progression by falling perfect fifth, such as in what we would call today a dominant resolving to a tonic, whereas UD successions appear as inflections of harmonically distant gestures, for example chromatic third relationships. This would echo Francès's observation mentioned in the introduction that expressive intonation in the same direction as a strongly expected resolution, such as a slightly sharpened leading tone resolving to a tonic, are less noticeable to trained listeners than would be a flattened leading tone in the same context. It could be that harmonic function and the effect of rhyme interact in our experiment, a possible avenue for future research.

## Conclusion

Evidence from the two experiments supports the hypothesis that musically trained listeners can generally detect the differences between 12-TET and 31-TET, in some cases even when steps of a fifth-tone are not involved. Discrimination performance is best for excerpts in which the tuning of the second chord of the excerpt differs greatly between 12-TET and 31-TET. The observed asymmetry in performance for the UD and DU groups of excerpts suggests that specific musical contexts may affect discrimination performance. The analysis of musical parameters supports this hypothesis; total displacement, absolute difference in distances on the Tonnetz, and the difference in cents between the second chords of the excerpts were all found to result in a significant increase in discrimination performance. Increased duration resulted in a decrease in performance, suggesting that memory limitations also played a role. The post-hoc analysis of specific pairs of excerpts also provided evidence that voice-leading patterns and directionality may overshadow discriminable acoustical differences. In some cases, the directionality of the voice-leading pattern can affect discrimination when the total differences in pitch and the voice-leading distance are held constant. Furthermore, excerpts sharing the same chords but different voice-leading elicited significantly different discrimination performance, yet one pair of UD and DU excerpts with the same voice-leading pattern but at different pitch levels did not elicit significantly different discrimination performance. We have thus demonstrated that specific voice-leading patterns and musical context can affect discrimination of the tuning systems. These results indicate that certain voice-leading

metrics may be more closely tied to perceptual phenomena than others. Theories of voice-leading that reflect perceptual tendencies in listeners can be of great value to analysts who wish to unite the mechanics of scored music with the audible experience of that music.

The results of Experiment 2 confirmed that participants relied heavily on the difference in pitch height between the final chords of the excerpts, but this ending effect was not symmetrical and could not wholly account for the differences in performance between the UD and DU groups. Future work along these lines is necessary to isolate and simplify the individual voices of the stimuli to target the psychological processes involved.

Through analytical and empirical study of Vicentino's use of steps of a fifth-tone, this paper has shown that the nuances of Vicentino's microtonal tuning system are generally (but not always) perceptible to trained musicians. This study provides preliminary information for further investigation into the perception of different musical tuning systems, the perception of various tuning systems in different styles of music, and the behavioural responses to these diverse, but often overlooked musical resources.

## Acknowledgments

This research was supported by a grant from the Fonds de Recherche Québec — Société et Culture (FRQSC) to JW (in collaboration with Ichiro Fujinaga and Peter Schubert), a grant from the Canadian Social Sciences and Humanities Research Council (SSHRC) to SM and JW, and SM's Canada Research Chair. We would like to thank Peter Schubert for his role in the production of the Vicentino recordings, Nicholas Squire for preparing the experimental stimuli from those recordings, and Bennett K. Smith for help with programming the experimental interface.

## References

- Barbour, J. M. (1951). *Tuning and temperament: A historical survey*. East Lansing, MI: Michigan State Press.
- Barker, A. (1984). *Greek musical writings I: The musician and his art*. New York: Cambridge University Press.
- Barker, A. (2007). *The science of harmonics in classical Greece*. New York: Cambridge University Press.
- Bent, M. (2002). *Composition, counterpoint, and Musica Ficta*. New York: Routledge.
- Berger, K. (1980). Theories of chromatic and enharmonic music in late 16<sup>th</sup> century Italy. Edited by George Buelow. *Studies in Musicology*, no. 10. Ann Arbor: UMI Research Press.
- Bigand, E. & Tillmann, B. (2005). Effect of context on the perception of pitch structures. In C. J. Plack, A. J. Oxenham & R. R. Fay (Eds.), *Pitch: Neural coding and perception*, (pp. 306-352). New York: Springer.
- Burns, E. M. (1999). Intervals, scales, and tuning. In D. Deutsch (Ed.), *The psychology of music*. 2<sup>nd</sup> ed., (pp. 215-264). San Diego: Academic Press.

- Burns, E. M. & Campbell, S. L. (1994). Frequency and frequency-resolution by possessors of absolute and relative pitch: Examples of categorical perception? *Journal of the Acoustical Society of America*, 96(5), 2704-19.
- Burns, E. M. & Ward, W. D. (1978). Categorical perception—phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America*, 63(2), 456-468.
- Cohn, R. (1998). Introduction to neo-Riemannian theory: A survey and a historical perspective. *Music Theory Spectrum*, 42(2), 167-180.
- Deutsch, D. (1973). Interference in memory between tones adjacent in the musical scale. *Journal of Experimental Psychology*, 100(2), 228-231.
- Galilei, V. (c. 1586). *Discorso intorno all'uso dell'Enharmonio*. Quoted in Palisca (1954).
- Hall, D. E. & Hess, J. T. (1968). Perception of musical interval tuning. *Music Perception*, 2(2), 166-195.
- Houtsma, A. J. M. (1968). Discrimination of frequency ratios. *Journal of the Acoustical Society of America*, 44, 383(A).
- Krumhansl, C. L. (1998). Perceived triad distance: Evidence supporting the psychological reality of neo-Riemannian transformations. *Journal of Music Theory*, 42(2), 265-281.
- Lowinsky, E. E. (1967). The musical avant-garde of the Renaissance or: The peril and profit of foresight. In C. S. Singleton (Ed.), *Art, science and history in the Renaissance*, (pp. 113-162). Baltimore: Johns Hopkins Press.
- Maniates, M. R. (1996). Introduction. In N. Vicentino, *Ancient music adapted to modern practice*, trans. Maria Rika Maniates, ed. Claude V. Palisca, (pp. xi-lxiii). New Haven: Yale University Press.
- McKinney, T. (2009). *Adrian Willaert and the theory of interval affect: The Musica Nova madrigals and the novel theories of Zarlino and Vicentino*. Burlington, VT: Ashgate.
- Miller, M. (2011). *Nicola Vicentino and the enharmonic diesis: An analytical and empirical study*. MA Thesis, McGill University.
- Miyazaki, K. (1993). Absolute pitch as an inability: Identification of musical intervals in a tonal context. *Music Perception*, 11(1), 55-72.
- Moore, B. C. J. & Moore, G. A. (2003). Discrimination of the fundamental frequency of complex tones with fixed and shifting spectral envelopes by normally hearing and hearing-impaired subjects. *Hearing Research*, 182, 153-163.
- Moran, H. & Pratt, C. C. (1926). Variability of judgments on musical intervals. *Journal of Experimental Psychology*, 9, 452-500.
- Murdock, J.B.B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, 64, 482-488.
- Palisca, C. V. (1954). *The beginnings of baroque music: Its roots in sixteenth century theory and polemics*. PhD thesis, Harvard University.
- Rakowski, A. (1990). Intonation variants in musical intervals in isolation and in musical contexts. *Psychology of Music*, 18(1), 60-72.
- Rakowski, A. & Miskiewicz, A. (1985). Deviations from equal temperament in tuning isolated musical intervals. *Acoustical Archives* 10(2), 95-104.
- Rogers, N. & Callender, C. (2006). Judgments of distance between trichords. In M. Baroni, A.R. Addessi, R. Caterina, & M. Costa (Eds.), *Proceedings of the 9<sup>th</sup> International Conference on Music Perception and Cognition, Bologna, Italy, August 2006*.
- Vicentino, N. (1555). *Ancient music adapted to modern practice*. Translated by Maria Rika Maniates. Edited by Claude V. Palisca. New Haven: Yale University Press, 1996.
- Vos, J. & van Vianen, B. G. (1985). Thresholds for discrimination between pure and tempered intervals: The relevance of nearly coinciding harmonics. *Journal of the Acoustical Society of America*, 77(1), 176-187.

- Ward, W. D. & Martin, D. W. (1961). Psychophysical comparison of just tuning and equal temperament in sequences of individual tones. *Journal of the Acoustical Society of America*, 33(5), 586-588.
- West, B., Welch, K. B. & Galecki, A. T. (2007). *Linear mixed models: A practical guide using statistical software*. Boca Raton: Chapman and Hall/CRC.
- Wild, J. (2014). Genus, species and mode in Vicentino's 31-tone compositional theory. *Music Theory Online*, 20(2). <http://www.mtosmt.org/issues/mto.14.20.2/mto.14.20.2.wild.php> (Last accessed, 1 July 2014).
- Wild, J. & Schubert, P. (2008). Historically informed retuning of polyphonic vocal performance. *Journal of Interdisciplinary Music Studies*, 2(1-2), 12

---

<sup>1</sup> In the system of genera, the three intervals contained within a given tetrachord are the only available intervals for that genus. The diatonic genus consists only of whole tones and a major semitone; the chromatic genera consist of the minor third and two sizes of semitones; and the enharmonic genera consist of the major third, the major enharmonic diesis (equal in size to the minor semitone) and the minor enharmonic diesis, which is a fifth of a tone in size in Vicentino's adapted genera. See Barker (2007) for more information about the genera of Ancient Greece.

<sup>2</sup> Vicentino confounds a number of tuning traditions in his treatise, and the result is unclear at best. For theoretical endeavours, he favours the mathematical rationality of the pure intervals; for practical purposes, he favours the irrational compromise of the geometric intervals of temperament. He even states in Book III on Music Practice that his enharmonic diesis is irrational and disproportioned, but that singers should be able to execute and accompany any size of interval to produce harmony. He does not, regrettably, entirely dispense with pure interval ratios in favour of temperament, and these systems are irreconcilable. Barbour (1951) surmises Vicentino intended a system of 31-tone quarter-comma meantone due to the tempering Vicentino describes. Wild (2014, cf. footnote 10 and paragraphs 7 and 8) discusses the extent of the difference between quarter-comma meantone and 31-tone equal and shows that they are moot in the case of Vicentino's compositional practice.

<sup>3</sup> Given that SAS does not allow post-hoc contrasts between excerpts in different groups when the model includes the effect of excerpt nested within excerpt group, these two contrasts were conducted within a model that includes only the effect of excerpt (without the nesting within excerpt groups), both as fixed and random effects.



## Biographies

**Stephen McAdams** studied music composition and theory at De Anza College in California before entering the realm of perceptual psychology (BSc, Psychology, McGill University, 1977; PhD, Hearing and Speech Sciences, Stanford University, 1984). In 1986, he founded the Music Perception and Cognition team at IRCAM in Paris. He was Research Scientist and then Senior Research Scientist in the French CNRS from 1989 to 2004. In 2004, he took up residence at McGill University as Professor and Canada Research Chair in Music Perception and Cognition in the Schulich School of Music.

**Mikaela Miller** completed her Bachelors degree in Music Theory and Composition at Arizona State University in 2008. She continued her studies at McGill University, where she graduated with a Masters degree in Music Theory in 2011. Her research interests in music include tuning systems, sensory perception, and mathematical models. She has presented at conferences at the University of Western Ontario and the University of Toronto, as well as at the 2011 meeting for the Society for Music Perception and Cognition. Ms. Miller is currently pursuing advanced degrees in Epidemiology and Biostatistics at the University of Colorado in Denver.

**Jonathan Wild** teaches music theory and analysis at McGill University. He holds a PhD from Harvard University (2007). His research interests include tuning, from both historical and speculative perspectives; analysis of late-tonal and early 20th-century chromatic repertoires; mathematical and computational methods for music theory; and corpus-based analytic methods. He is also a sought-after composer of vocal music and has received over 50 performances of his works by the Hilliard Ensemble.

**Bruno L. Giordano** holds a PhD from the University of Padova, Italy jointly supervised by STMS-IRCAM-CNRS, Paris, France (2005). His research focuses on the perception and cerebral processing of non-speech natural sounds, and on the psychophysics of audio-haptic integration. In 2011 he was awarded a prestigious Marie Curie fellowship (FP7 PEOPLE-2011-IEF-30153) to carry out his brain-imaging research at the Institute of Neuroscience and Psychology of the University of Glasgow.