Historically Informed Retuning of Polyphonic Vocal Performance

Jonathan Wild and Peter Schubert
Schulich School of Music, McGill University

Background in history of music theory. The use of just intonation has generated heated debates ever since the first accounts of the articulation of pitch space in the scientific treatises of Ancient Greece. In the Renaissance, the discussion turned to vocal music as the locus of debate. One of the problems with using just intonation is the incommensurability of pure intervals with one another (e.g., the A four pure fifths above F is not the same pitch as the pure major third above F). Treatises and accounts of tuning systems proliferated, and our present-day understanding of what is at stake is limited by the dearth of sounding examples.

Background in performance. On the one hand it is very difficult to verify precisely how modern-day singers tune, and on the other, although recent interest in historical tuning has generated several articles with electronically produced sound examples, the examples do not contribute to a direct understanding of tuning in the vocal context—synthesized sound is simply not “the real thing.”

Aims. Our study aims to show how the gap between theories of tuning and the art of vocal performance may be bridged.

Main contribution. By producing recordings of precisely tuned audio in rich vocal timbres (with vibrato, consonants and vowels), we make available actual repertoire examples, reliably tuned, in a human-sounding rendition.

Implications. Our recordings will provide music scholars with the sensory data that has been missing from accounts of various speculative tuning systems. In addition, it will provide performers with convincing models on which to model their own tuning in performance.

Keywords: Tuning, just intonation, Benedetti, Zarlino, Lassus, vocal performance, Melodyne

*Correspondence: Jon Wild, Schulich School of Music, McGill University; e-mail: wild@music.mcgill.ca.
Introduction

In the following pages we examine some of the issues inherent in the use of just intonation, as applied especially to vocal music in the light of Renaissance debates on the topic. A principal contribution of the current research is a means of rendering convincingly retuned versions of musical examples from repertoire and from the theoretical accounts of previous centuries. We accomplish this with simple recording techniques and commercially available post-production software, making it possible for the modern listener to hear a vocal group singing according to performance practices theorized in the 1500s. Audio examples for this article may be accessed through the website of the Journal of Interdisciplinary Music Studies.

Background

From the point of view of performance practice, the last few decades have seen a burst of interest in historical tunings on the part of early music instrumentalists, especially keyboardists. But singers have not been as quick to respond rigorously to the expressive possibilities of alternative tuning systems—their instrument has no fixed points of reference like frets, keys or valves; instead it provides a flexible pitch continuum, with all the possibilities and difficulties that that implies. And we have found that even when early music vocal groups have a sophisticated awareness of the issues involved, and have the best intentions, the practical results cannot necessarily be relied upon—in other words their intonational practice might be accomplished and tasteful, but the group may not tune precisely how they claim to. Compounding the difficulty is that it is a non-trivial signal-processing problem to obtain precise intonational measurements for the individual voices in a polyphonic recording; the issue is described more thoroughly in another contribution to this journal (Devaney & Ellis 2008). Thus the first problem generally encountered in performing alternatively tuned renditions is that of pitch accuracy; our solution allows independent control over the tuning of each voice to a high degree of accuracy.

From a theoretical point of view, several articles on historical tuning systems have been published in online journals in recent years; these have been able to include aural examples, produced and tuned with great precision, by synthesisers. The timbres are not always entirely pleasant to listen to; their unnaturalness can distract attention from the tuning. Our techniques produce aural examples in natural vocal timbre, since we begin with recordings of real singers and perturb the perceived vocal qualities as little as possible. Thus for the first time it is possible to solve simultaneously both the problem of pitch accuracy and the problem of timbre, granting the listener access to performances in previously unrealisable tunings. While the techniques described are applicable to a broad range of music composed for non-standard tunings we concentrate here on basic historical tuning issues for the uninitiated and on our technical methodology.
History of Theory

Arguments about the division of musical space formed a significant portion of the surviving body of ancient Greek music theory; certainly it is true to say that tuning has been a vexed issue ever since. The Greeks approached the question from a purely melodic point of view, but the debate becomes especially thorny when we consider harmonic music—that is, music that has a vertical, simultaneous dimension as well as melodic. It became apparent to the theorists of the Renaissance that there was a problem inherent in its emerging triadic music, a problem that stemmed from the essential incommensurability of the consonances of the system: octaves, fifths and thirds. These harmonic intervals sound best—so the prevailing thought went—when they are tuned acoustically pure, that is, when the intervals as sung match the intervals found in the overtone series, which have small integer ratios. To greatly approximate a complex situation, we can say that when the frequencies of two pitches are in a simple ratio, the resulting interval will sound more consonant than when they are not in a simple ratio. The main problem with using these pure intervals as not only the model for perceptual preference, but also as the basis for a musical system, is that they are essentially incommensurable with one another: when you stack one of them—that is, you repeat it or add it to itself—you can never produce an exact multiple of any of the other intervals. For instance, stacking perfect fifths—whose just ratio is 3:2—will never produce a just major third (5:4), or a major third compounded by any number of octaves (2:1), or indeed any multiple of a major third. This is clear mathematically: \((3/2)^n\), or \(n\) fifths stacked, will have a numerator that is a power of 3—which by the fundamental theorem of arithmetic (the uniqueness of prime factorisations) cannot equal the numerator of a stack of major thirds, \((5/4)^m\).

This incommensurability can be conveniently visualized on a harmonic lattice known to music theorists as the Tonnetz, associated with the nineteenth-century theorist Hugo Riemann, which shows fifth-relations and third-relations. Figure 1 depicts a small portion of the Tonnetz, which must be conceived as indefinitely extended in both dimensions of the plane. Relationships of a perfect fifth are shown as connections on the horizontal axis; if we ignore octaves (i.e. casting out powers of two) then motion rightwards on this horizontal axis represents a tripling of the frequency, and leftwards a division by 3. Motion in either direction along the rightwards-ascending diagonal axis, where pitches are related by major thirds, represents multiplication or division of the frequency by 5. These two representations constrain the remaining axis, rightwards-descending, to represent just minor thirds, or the difference between perfect fifths and major thirds. Thus (still disregarding powers of two, which correspond to octave shifts) the Tonnetz interpreted as frequencies relative to a central 1:1 would appear as Figure 2. It should be apparent that radiating outwards from the central 1:1, all ratios involving 3s and 5s in any combination will eventually appear.
Figure 1. The Tonnetz, a harmonic lattice of fifths and thirds.

Figure 2. The Tonnetz as a lattice of purely tuned frequency ratios.
When the intervallic relationships are taken to be pure, a given note-name appearing in two different rows of Figure 1 will represent slightly different pitches. So for example, stacking four perfect fifths above an F will produce an A that is different from the A found a pure major third above the same F. Riemann used over- and under-scores with the note names to show this difference on his Tonnetz, but here we have called both pitches simply “A”, keeping in mind that A’s on different rows of the Tonnetz refer to different pitches. The difference between pitches on adjacent rows with the same note name is known as the syntonic comma and has the ratio 81:80, working out to approximately 21.5 cents.

The first major figure associated with the claim that singers used pure intervals was the Renaissance theorist Zarlino, who in 1558 put forward in his treatise *Le istitutioni harmoniche* a tuning that he called the syntonic diatonic, after a scale described by Ptolemy, the ancient Greek astronomer and scientist. We tend to call this scale the just intonation scale, though in reality it is one of many possible just intonation scales. Figure 3 shows Zarlino’s graphic representation of the relationships between notes in a C-to-c major scale—every arc stands for a purely tuned interval. On the portion of the lattice diagram that appears in Figure 4, we can show all of these same relationships in a visually simpler format. It follows from the triangles connecting the notes of the major triads on each of F, C and G that these are tuned purely; the F major triad is highlighted on the diagram. Pure minor triads will appear as connected inverted triangles; Zarlino’s scale gives us a pure minor triad on each of A and E (the triad on E is highlighted; it appears as an inverted triangle). The glaring problem when the notes are tuned in this way is that there is no triad on D: the 3rd (F) and the 5th (A) are off in a part of the lattice removed from the root. Without a direct lattice connection as fifth and third the pitches shown do not form a pure triad with the root. Attempting to form a D-minor triad from these three pitches results in an unpleasant and mistuned sonority, often called a “wolf” of the system. When confronted with this unavoidable defect, Zarlino said that still, given the flexibility of the voice, singers would be able to adjust so that a triad on D came out pure anyway—but this is really tantamount to confessing that they do not sing using only the notes of the syntonic diatonic after all.
Figure 3. Zarlino’s harmonic division of the octave, the *syntonic diatonic* (Zarlino 1558/1965: 122).

Figure 4. Syntonic diatonic as a portion of the Tonnetz, with F-major and E-minor triads highlighted with superimposed triangles.
Benedetti and syntonic comma drift

Our first musical example is due to Giovanni Benedetti, an Italian scientist of the 16th century. He composed this example in a letter to a friend in 1563, pointing out that if singers performed it using pure intervals, the music would end on a different pitch from its beginning. Benedetti’s passage takes the inherent difficulty in using strict just intonation—the fact that the incommensurability of the different consonances in their pure forms leads to drifts in pitch—and distills it into the simplest example he could construct. He did not say whether he thought the resultant drift would be a bad performance outcome or not; he used these examples to contrast the behavior of voices with that of instruments with a fixed gamut of pitches.

The music in Example 1 consists of a two-bar pattern that repeats four times, and we map just one iteration of that pattern onto the partial lattice of Figure 5a. If we let the voices tune so that the pitch of any common tones between chords in the progression is preserved exactly, and any new notes are tuned pure with respect to the pitches preserved as common-tones, we will find that over the course of the two bars the tonic G drifts by one row in the lattice—that is, by one syntonic comma. The numbers 1 to 5 above the score of Example 1 correspond to the locations labelled on the lattice of Figure 5a. If we repeat this music four times as in Benedetti’s example, the amount of drift will accumulate accordingly. The gradual rise in pitch of Audio Example 1a is more noticeable to some listeners than others; it is unmistakable when we isolate the first and last chords as in Audio Example 1b.

Example 1. Passage by Giovanni Benedetti.

Figure 5a. Portion of the Tonnetz showing drift in the passage by Benedetti.

Benedetti’s example has previously been discussed by Claude Palisca (who believes Benedetti’s intention was to show Zarlino was wrong to champion just intonation).
and a synthesized version has previously been realized by Ross Duffin. The present version in natural human timbre permits a more musical judgment on the part of the listener, and provides a natural model for singers to emulate.

The drift in pitch is not a very satisfactory solution—whatever Benedetti believed, other Renaissance writers knew that singers did not actually do this and that in practice, just intonation would not be used so strictly. A different solution, which permits the use of purely tuned chords and yet manages to avoid the drift, is to use two slightly different versions of a given note. In the alternative rendering we have made of the same music in Audio Example 1c and diagrammed in the lattice of Figure 5b, we have shifted the pitch A in the soprano; you will first hear the A tuned two pure fifths above G, to go with the D; then the A tuned a minor third below the C that is subdominant to G. The second A that the soprano sings is a syntonic comma lower than the first. And that second A leads to an E and C that are in the correct place with respect to the original G, so that each repetition begins on the same G; thus on the lattice diagram of Figure 5b the numbered labels “1” and “5” are at the same location. We have isolated the soprano part in Audio Example 1d, to help in focusing on her melodic shift. The two sizes of whole tone between her G and A are 9:8 and 10:9, which are different by one syntonic comma or ~21.5 cents (just over a fifth of an equal-tempered semitone). This approach works, but requires extreme pitch sensitivity and control on the part of the singer charged with effecting the syntonic shift.

In his example, Benedetti wrote one A, tied across the barline; however, in order to facilitate the pitch change, we have a rearticulated note. With the tied A, the soprano’s inflection would result in a little glissando during one sustained note, rather than two discrete notes of different pitches. We anticipated it would be easier to do the retuning with two different notes, so we had the soprano rearticulate every half note. But this suggests the possibility of other kinds of “solution” to the problem posed by Benedetti’s passage, such as attacking the A on the downbeat at the common-tone pitch that feels appropriate melodically—even if that puts it high with respect to the C and E that are adjacent to the G on the lattice—then sliding down a syntonic comma to settle on a pure consonance with the C and E. If performance practice permits sliding like this during the sustained portion of a note, many problems of the syntonic comma evaporate, as the melodically convincing pitches at the attack of a syllable
may always subsequently glide to the pitches that are required for a purely intoned sustained consonance.

Example 1, as we mentioned, was an artificially simple passage, but the same kind of progression occurs all the time—both forwards and backwards—in actual music. Example 2 shows a slightly more fleshed-out version in four parts; as in the Benedetti the pitch of the final triad is one comma sharper than the beginning triad. The short passage with its drift of a syntonic comma sounds fairly innocuous (Audio Example 2), but in longer pieces these discrepancies can accumulate: a syntonic comma here and a syntonic comma there, and pretty soon you can have a very noticeable offset!

Example 2. Four-part progression exhibiting comma drift.

Other sizes of drift

The harmonic lattice of Figure 1 may be used to map drifts of other sizes than the syntonic comma, too. Example 3, Figure 6a and Audio Example 3 illustrate a drift of a lesser enharmonic diesis, produced by moving the root through three successive major thirds. This passage, when tuned preserving common tones, ends on a triad 41 cents sharper than the one it began on. In terms of the frequency calculations, $(5/4)^3 \approx 2$.

Example 3. Cycle of major thirds.

Example 4, Figure 6b and Audio Example 4 illustrate a drift of a greater enharmonic diesis, produced by moving the root through four successive minor thirds. This passage ends 63 cents flatter than it began. The frequency relationship is $(6/5)^4 \approx 2$. These last two examples show that in just intonation, notes that we usually think of as equivalent (in the familiar 12-tone equal temperament), like A♭ and G♯, will not only have different tunings, but also that the difference in pitch between them will depend on the path taken to connect them.\textsuperscript{10}
Example 4. Cycle of minor thirds.

Figure 6. **a)** Lattice diagram for cycle of major thirds, **b)** Lattice diagram for cycle of minor thirds.
Pure tuning applied to an entire composition: Lassus’s *Prophetiae Sibyllarum*

The Tonnetz is also a useful visual aid for following harmonic motion in a longer piece of real repertoire. Example 5 shows one of the famously chromatic motets from Lassus’s cycle *Prophetiae Sibyllarum*, from about 1558. “Famously chromatic” in this case does not preclude a high degree of consonance—the composition consists basically of homophonic triads, almost all major—rather, it is the relationships between juxtaposed triads that are very often unexpected. This makes it an interesting piece to consider how to tune.11 Given all the root-position major triads, the composition is an especially convincing candidate for an application of pure tuning. The principles we have used for the first tuning (Audio Example 5) are to keep all vertical sonorities pure; to keep all common tones between successive chords the same (i.e. no melodic shifts); and in cases where there are no common tones, we chose the tuning that best preserves the tonal context on the short scale. Figure 7a shows the tuning of the first page of music; the harmonic progression is indicated by numbers in the lattice corresponding to the labels of Example 5. After only nine measures we have already drifted upwards by two rows on the lattice; by the end of the motet, tuning according to this model, there is a cumulative drift of five rows, representing a flattening of pitch of more than a semitone relative to the beginning. The fact the drift is consistently in this direction—upwards on the lattice; downwards in pitch—has interesting implications for Lassus’s harmonic syntax.

Whenever there is stepwise root motion, there are no common tones to anchor the new chord, so there is a choice on the lattice. Going from B major to C♯ minor in bars 3-4, for example, the root can move either rightwards by two positions on the lattice (producing a 9:8 whole tone between B and C♯) or leftwards by one position and left-and-up by another (producing a 10:9 whole tone between B and C♯). Making different choices at the junctures where there are no anchoring common tones will result in different drifts over the course of the piece. We think that because of the context of the B-major triad (labelled “3” in the score) moving to a C♯-minor triad (labelled “4”)—that is, the context of a slightly larger-scale move to an E-major triad (labelled “5”)—the note B shared by the B major and E major should be tuned the same way. This is what we mean by choosing the tuning that best preserves the tonal context on the short scale.
As it turns out, this preference led to a greater amount of drift over the course of the whole piece. Moving instead to the C♯-minor triad indicated by the label “4” on Figure 7b will “save” one of the comma shifts that occurred in Figure 7a; a similar choice when E major (“5”) moves to F♯ minor (“6”) saves the other comma shift, and the progression as tuned in Figure 7b (and Audio Example 6) returns to its starting point. These tuning choices cannot be made by the performers as the piece unfolds without foreknowledge of the future progression of the music; there is no reason to move roots (in the bass voice on both occasions) by 10:9 whole tones instead of by 9:8 whole tones, except if we know that we will need to compensate for a drift upwards on the lattice (downwards in pitch).

Figure 7b (Audio Example 6) took advantage of both root progressions by step to correct the overall pitch drift; in so doing, however, the music visited a tonic triad (G major, labelled “8”) that was not the same as the beginning and ending tonics. An alternative solution is shown in Figure 7c and realised in Audio Example 7. Here the first stepwise shift, “3”-“4” is accomplished the same way as in Figure 7b, but proceeds similarly to Figure 7a afterwards. In order to finish at the right pitch level, a melodic shift of a syntonic comma must be introduced in the alto voice between the B♯-major (“11”) and D-major (“12”) triads, rather than preserving the common tone D at pitch. Even for this tuning solution which stays closely around the “home” tonic on the lattice, no fewer than seventeen pitch-classes are required, including two each of D, A, E and B, and both members of the enharmonic pair D♯/E♭.
It should be obvious from the preceding brief discussion that settling on the best tuning for a passage of music like this is no simple matter. Our retuning methodology allows us to evaluate the different solutions, to illuminate and bring to life the theoretical debates of the past. Eventually the method could be used to train singers to reproduce the subtle differences between tuning systems in their live performances; we have found it helpful, in exploring the vocal practices required in various tuning solutions, to generate “music minus one” versions of compositions, with one vocal part omitted from the mix. The missing part may be performed live by one of the singers; once they are able to perform it reliably, subsequent stages in the training process for a vocal group could include “music minus two”, etc., until finally all the members have been brought in.

**Retuning Methodology**

*Melodyne* is commercially available software created by the German company Celemony; its principal developer is Peter Neubäcker. Melodyne has made quick inroads in the field of popular music post-production, especially tuning “correction”, thanks to its powerful and flexible processing possibilities controlled by a visual interface that appeals directly to the user’s musical intuitions. While the software has no doubt found the vast majority of its applications to be in contexts where equal temperament was the goal, it is possible to adapt it to user-defined tuning systems. The adjustments of tuning and timing are accomplished with minimal disruption to sound quality; notable for vocal research is the fact spectral formants are preserved throughout audio transformations. (Melodyne’s analysis/resynthesis engine is proprietary; in any case the precise details of its inner workings are not germane to the present discussion.)

An essential requirement for working on polyphonic music with Melodyne is that each part be isolated in its own monophonic track (“monophonic” here refers to a sequence of musical pitches that occur only one at a time; recordings may be stereophonic and still work). Naturally this poses a problem; existing recording techniques that attempt to isolate musicians performing simultaneously in the same
Historically Informed Retuning of Polyphonic Vocal Performance

Acoustic space (e.g. closely focussed chest microphones) were rejected as not providing sufficient acoustic isolation, which would lead to artifacts in the reengineered recordings. Our solution was to make a “normal” recording of the vocal group singing together—the “reference” recording—then to bring each member of the group into the recording studio separately and record them singing—the “solo” recording—while they listened to the reference performance played back through headphones. (Many singers preferred to use only one earpiece in this situation so they could hear themselves well enough. Amount of emphasis on the singer’s own part in the headphones was controlled by having the reference recording made with multiple microphones recording to separate tracks. Thus, for example, the alto could choose to sing along to the reference recording mixed so that her own part was relatively attenuated compared to the others if she wanted; in fact this was generally preferred by the singers. There was always some sound from their own part present, though, because of bleed-through from the other microphones.) Our solution allowed us to achieve good temporal coordination between parts, and an overall agreement on dynamic shaping, since each singer was responding to the same reference recording. The reference recording was made with the singers tuning as came naturally to them, rather than attempting to produce themselves the micro-intonational shifts we were investigating. Some of the drawbacks of the method will be detailed below.

Once the individual parts have been imported into Melodyne, the editing can begin. Melodyne presents notes visually as “blobs” on the screen that can be selected, dragged, and otherwise edited. While the present article cannot provide a complete tutorial on using the software, the following points outline our working method for the benefit of researchers hoping to achieve similar results.

1) Upon examination of the coordination between solo recordings, any problems of timing were first dealt with by moving onsets and extending or compressing durations. In many cases, extending a note by more than a few percent resulted in audible problems with the attack portion, and so the noisy attacks of many sung syllables were separated out and extension/compression applied only to sustained portions of the note.

2) Melodyne automatically detects note boundaries with quite good accuracy, but sometimes treats one note as two, or two notes as one. These problem spots had to be corrected manually.

3) While Melodyne can adjust the average frequency of a given note to within one cent’s accuracy, it allows the user to preserve the recorded deviations in pitch from that average if desired. Thus the whole of the captured audio, including vibrato, may be transported to a new pitch. We asked the singers to sing relatively “straight”, i.e. with little vibrato. This was to enhance the “locked-in” percept associated with just intonation chords. However not all singers were equally adept at straight-tone singing, and so we applied some measure of vibrato attenuation in those cases. This is a very tricky signal-processing problem, as the cyclic modulation of the vocal output in some singers’ vibrato occurs not only in the
realm of frequency, but also in amplitude and in spectral content. In other words some of the vibrato is “deeper” in the sound than mere variation in pitch. So in some cases, reducing the naturally-occurring pitch variation by a sufficient amount to bring out the justness of the tuning led to a certain unnatural vocal quality. Until such time as the signal-processing problem has been solved more adequately, future recordings should either use singers with excellent pitch stability (this was the case for three of our four singers, and so should not be unreasonably hard to attain), or should allow vibrato to remain at its original levels.

The attenuation of pitch modulation (vibrato) should not be applied uniformly throughout the duration of each note when a noisy attack is present, especially for sibilant phonemes like “ss”. Any reduction in the variation in pitch for these portions of a speech signal will result in a noticeable “robotic” quality to the sound. It is therefore necessary to separate out the attack portion of the note and leave it unmodified, to preserve the natural quality of the edited sound.

4) The score must be analysed and the intended pitch of each note (in the context of the tuning system under consideration) determined, relative to equal temperament. We make a note on the score of each pitch’s intended offset from equal temperament, in cents. This information will be used to retune the pitches in Melodyne’s arrangement of the recording. Conceptually the simplest next step is to move each note manually to the desired pitch, though a faster method will be described in the following paragraph. If any notes have been separated into distinct parts during the editing operations already undertaken (e.g., attenuating vibrato), it will be important to move all portions of the note together as one.

In addition to allowing manual retuning of each note individually, Melodyne permits the user to specify a scale that it will retune the music to automatically (as long as that scale has no more than twelve pitch classes). For the sake of simplicity, I will first consider the case where the intended retuning of the music has no more than twelve different pitches per octave. Before setting the scale, it is best to determine any overall discrepancy from $A = 440$ Hz, so that the entire scale is defined relative to some pitch that accurately represents a reference note as it sounds in the performance. Once the scale has been specified, by fixing values for each of the 12 pitch-classes, it is possible to select all notes in a track and have them “snap” to the nearest note of the defined scale. In some instances this is still not the intended pitch, for instance in a case where there are two variants of a D in the score and therefore two variants of D in the defined scale—call them $D_1$ and $D_2$, the former lower than the latter—and the singer performs a pitch sharper than either when the intended pitch was $D_1$. In such a case manual correction will be necessary. (The scale may contain two Ds as described here, but one of them will have to be (mis)labelled as C# or D# in the scale definition box, because Melodyne’s labels for the twelve pitch-classes are fixed and cannot be changed.)

More often than not, the gamut of notes required to tune a composition in just intonation will involve more than twelve pitch-classes. The limitations of
Historically Informed Retuning of Polyphonic Vocal Performance

Melodyne means that for now it is necessary to accomplish the retuning in multiple “passes”, each retuning some subset of the pitches involved, never more than twelve at a time. Separate scales must be defined that taken together exhaust the pitches required to tune the composition. Intelligent partitioning of the pitches involved, into scales whose notes are sufficiently separated from one another, will make this stage of the process run more smoothly, as fewer manual corrections will be necessary. A further limitation of Melodyne is that this operation must be performed on each vocal track separately, at least with the current version of the software.

5) Once the individual tracks have been edited, they must be mixed together and balanced properly for a convincing final performance. Melodyne has an integrated mixer, but for higher quality we imported the tracks into a ProTools session and did the mixing there. We also added some reverb, because the solo recordings had to be undertaken in the acoustically dry environment of the studio—no note was allowed to sound together with the decay of the note preceding it, or Melodyne’s pitch correction algorithms would have been confused. (The reference recording was done in a reverberant environment in anticipation of this stage of the process, since performers respond to reverb in their tempo and timing.)

Another possible method for achieving similar results, replacing the two-pass process of reference recording and solo recording, would be to use glass isolation booths for recording—so the performers can see each other for visual cues—and to pipe in the audio from the other singers through headphones. A conductor, visible to all involved, could also participate if necessary.

We have not yet attempted to retune recordings of choirs (i.e. multiple voices on each notated part). Possible approaches in that scenario would include recording each “section” of the choir separately after the reference recording—this would require musically “tight” sectional performances, to avoid the problem of multiple pitches sounding simultaneously in the individual track—or, if that proved unfeasible, recording each singer separately, which would result in a proliferation of solo tracks and a potentially greater number of synchronisation difficulties.

One weakness we perceive in our current approach is that the lack of spontaneity in the double recording process leads to artistic results that may lack drama; a possible explanation of this is that it is much less natural for singers to respond “with feeling” to a swell (for example) in music that is piped in to them through one earphone, than to music which surrounds them in live performance. Another weakness is that the singers cannot exploit the nuances and colouring of the tuning system in an interactive way, nor interact in their recorded performance with a genuine acoustic space. Despite these drawbacks, we believe a similar combination of natural vocal timbre and reliably accurate tuning has not been previously achieved.
References


For example, Rahn 1998; Wibberley 2004a, 2004b; Woodley 2006; and Duffin 2006.

2 For an account of Greek codifications of pitch space, see Barker 1989 and 2007.

3 Before the frequency-based model of pitch was discovered, these ratios were taken to be ratios of string-length rather than frequency—but the ratios themselves are the same in either case, because frequency is inversely proportional to string length. The extent to which consonance is actually founded in simple numerical ratios of frequency is explored in a great many publications in the field of psychoacoustics.

4 Kopp 2002 contains a good history of the Tonnetz.

5 Gioseffo Zarlino, 1558/1965: 122.

6 Audio examples for this article may be accessed through the website www.musicstudies.org.

7 Benedetti 1585/1924-25.

8 Most experienced listeners will recognise that the C when it enters in the bass, tuned to the already sounding A, is high. The bass (according to Benedetti) must sing a 10:9 whole tone instead of 9:8 here, in order to make a pure major sixth with the A—but our memory of the G and D is strong enough (again, speaking of experienced listeners) that we can tell something is wrong with the C.
Historically Informed Retuning of Polyphonic Vocal Performance

9 (Palisca 1985: 257-265; Duffin 2006). Duffin (para. 31 and Example 1c) uses a notation based on that of Easley Blackwood, where notes receive a superscript denoting any shifts of a syntonic comma relative to a Pythagorean chain of fifths. For example, a purely tuned major triad on G would include the pitches G0, B-1, D0 (or G1, B0, D1, etc.). This is entirely equivalent to a Tonnetz representation: notes with a superscript of “1” correspond to notes one row beneath the row of notes with a superscript of “0”; a superscript of “-1” means a note is from one row higher than the “0” row, etc.

10 We do not consider here perhaps the most famous comma, the Pythagorean comma, which in fact is not that important in this music—it is very rare that a complete cycle of fifths occurs or is implied.

11 Henry Klumpenhouwer (1992) offers a tuning solution to this piece in his article. His solution follows a model suggested by Descartes wherein each voice separately must be related by common tone, fifth or third to some voice in the preceding chord. Whenever root motion is by step, this algorithm produces an impure triad. And, as does our approach, his causes the piece to end at a different pitch level from its beginning.

12 The accuracy of one cent is more than sufficient for our purposes; Hagerman and Sundberg (1980) find barbershop groups operating within an average accuracy of 2 to 3 cents.