Towards a Difference Limen of Musical Scale Perception

Jim Russell – 21harmony.com

ABSTRACT

Every twelve-tone musical tuning has a quantifiable deviation in its step sizes from those of *Equal Temperament*. While larger differences are easily detectable to the human ear, a smaller difference may be unnoticeable.

Previous research into the Just Noticeable Differences (JND) or Difference Limen between concurrent tones of different frequencies provides a good estimate of the thresholds of human pitch perception under strict (laboratory) conditions and using simple signals. In a practical context such as music listening, these data are made less useful by the introduction of many simultaneous signals presented together.

A pilot study is conducted in which listeners are presented with a musical piece, the tuning of which drifts over the course of 90 seconds to one of three extents – 4.76 cents RMS (*Equal Temperament* drifting to *Meantone Temperament*), 10.15 cents RMS (*Equal Temperament* drifting to *Just Intonation*) and 0 (no drift in tuning – *Equal Temperament* throughout).

Listeners are more likely to report a drift in the music's tuning if one occurs, but are not seen to be statistically more likely to report one when presented with a stronger level of drift in tuning. Furthermore, the proportion of listeners reporting the two drifting tunings to have indeed 'gone out of tune' is approximately half. A better investigative model is suggested to inform further study.

A. INTRODUCTION

Background

Western music is currently overwhelmingly recorded, performed and heard in *Equal Temperament*, a system of tuning the 12 notes of the musical scale equally so that each note is related to the next by a frequency difference of 1.059 (the twelfth root of 2). Historical alternatives to this tuning system (such as *Pythagorean* tuning, *Meantone* and *Well* temperaments) are seldom encountered in popular music, and nor are other more contemporarily popular 'alternative' tuning systems such as *Just Intonation,* the *Bohlen-Pierce* scale, and various divisions of the octave into other than 12 parts.(1) (2) (3)

Recent times have seen a growing number of lively online social outlets(4) (5) (6) in which people can create and share music in these 'alternative' tunings. Given that a scale can be created with infinite resolution, having any number of steps placed at any particular point, the potential number of tunings is huge if not arguably infinite. This paper, though, will only be concerned with 12-tone tunings.

Rationale

In one study(7), the author compared approximately 180 historical tunings from (Barbour, 2004)(8) to detect 'cyclic acoustic matches'. These are two tunings that display the same contour in their notes' deviation from equally tempered, 100-cent steps.

Kepler's Monochord (cents)	Contour	Malcolm's Monochord (cents)	Contour
92	-8	112	+12
204	+12	204	-8
316	+12	316	+12
386	**-30**	386	-30
498	+12	498	+12
590	-8	590	-8
702	+12	702	+12
814	+12	814	+12
906	-8	884	**-30**
1018	+12	996	+12
1088	-30	1088	-8
1200	+12	1200	+12

Two acoustically matching tunings from Barbour, 2004 and their contours (the deviation from 100-cent ET steps from one note to the next)

The table shows the tunings of Just Intonation monochords by Kepler and Malcolm in cents (hundredths of an equally tempered semitone). These tunings are acoustically identical but start from a different position, 5 (or 7) steps removed. When music of the same key is played in both tunings, they will sound different, but the *contour* of each tuning is identical, only offset in this case. See (7) for a more detailed explanation.

This method makes a start in identifying any two tunings that are mathematically/acoustically identical using the cent as its smallest unit of resolution. This is a relatively arbitrary unit to deem two things identical, and while very accurate results of exact matches are reached, these may not in fact be the only *perceptual* matches. Slight variations in the contour of two tunings could be negligible to the auditory system's discrimination of them.

What, then, is the smallest unit of resolution in the contour of a musical scale that will lead to the listener detecting a difference?

Previous Research

Data from several researchers, most notably Fastl [(9)] and Moore [(10) (11) etc], explore the 'Just Noticeable Differences' (JNDs) or 'Difference Limen' in auditory perception. These JND 'thresholds of difference' have been investigated for basic factors such as amplitude changes, amplitude modulation, frequency changes, frequency modulation, phase, and the effect of different types of masking signals on each.

Based on the most relevant findings of the above sources, the JND value for frequency changes (Δf) is roughly 3 – 4 cents when the frequency (f) is above 500Hz, and increases below this value. Additionally, broadband noise masking the audio signals creates a doubling of the necessary difference if the tone and the noise are at the same intensity. Other tones in a musical melody may or may not act to influence perception thresholds in a manner akin to the broadband noises used in previous studies.

It would be inappropriate, however, to simply extrapolate the data from tests that used simple sine tones in order to reach even an estimate for a Difference Limen of scale perception in a musical context. Far more factors are at work in a real musical piece (complex tones, inharmonicity, varying rates of masking of tones by others, note intensities, durations of notes and of spaces between them to name but a few), as well as psychological factors such as pitch memory, musicality, training and experience.

Aims of the study

This pilot study will attempt to analyze listeners' perception of pieces of music that drift in tuning. By making the investigation relatively general and broad-ranging, it can make some first minor inroads into human perception of more complex stimuli in order to inform further studies that can 'pin down' how various musical factors contribute to these perceptual thresholds.

B. METHOD

To obtain a larger sample, the test was carried out using the internet. Respondents were presented with one of a possible three versions of a musical piece, two of which drifted out of their initial tunings, and one of which remained constant throughout. Respondents were then asked whether they "perceived the music to go out of tune (from start to finish)."

Musical Properties

The music used as a test stimulus consisted of two harps playing a counterpointed musical phrase that repeated once, lasting 1 minute 37 seconds in total. There were two incidents of chromatic passing notes in each repeated phrase. The notes used never went below the second or above the fourth octaves.

Audio Properties

The music was created in software using a sampler to gain more control over its consistency. The two harp instruments were presented one per stereo channel. It is noteworthy that the harp is a plucked string instrument – it is therefore subject to inharmonicity, but a good representation of the predominant timbral characteristic of recorded music.

Using Melodyne Editor software, the piece was retuned to *Equal Temperament* before the test pieces were retuned. (Though the piece had been programmed in *ET*, Melodyne's algorithm takes into account any inharmonic elements of the analysed sound, so therefore assigns certain notes slightly incorrect positive values, see [(12)]).

All three musical pieces started tuned to *Equal Temperament*. The notes of the two test pieces drifted to the destination tunings at the rate shown in the below tables, while the notes of the control piece remained constant throughout.

TEST PIECE 1
Destination Tuning: *Meantone Temperament*
Total drift by end (RMS): 4.76 cents
Range of drift: 12 cents

Time (secs)	**0**	**24**	**30**	**48**	**60**	**72**	**End**
Bb	0						**0**
C	0	+1		+2		+3	**+4**
D	0	-1		-2		-3	**-4**
Eb	0			-1			**-2**
F	0			+1			**+2**
G	0						**0**
A	0	-2		-4		-6	**-8**
(Db)	0	-1		-4		-6	**-8**
(Gb)	0		-2		-4		**-6**

TEST PIECE 2
Destination Tuning: *Just Intonation*
Total drift by end (RMS): 10.15 cents
Range of drift: 32 cents

Time (secs)	**0**	**24**	**30**	**48**	**60**	**72**	**End**
Bb	0						**0**
C	0	+1		+2		+3	**+4**
D	0	-4		-7		-11	**-14**
Eb	0			-1			**-2**
F	0			+1			**+2**
G	0	-4		-8		-12	**-16**
A	0	-3		-6		-9	**-12**
(Db)	0	+4		+8		+12	**+16**
(Gb)	0	+4		+7		+11	**+14**

Listener Conditions
At the benefit of sample size, no limits were placed on the method of listening. For a simpler test, listeners were able to listen to the music in whatever way they preferred, and were not instructed to do otherwise. Varying listening methods may bring more variance to the results, but allow listeners to respond using the methods with which they feel most comfortable and familiar. While the level of loudness and detail experienced by different respondents will have been largely variable, different listening methods shouldn't influence the experienced note frequencies to a detrimental extent.

C. RESULTS

From a total of 307 respondents:

Drift	Drift Reported	No Drift Reported
None (control)	**30**	**61**
RMS 4.76 cents	**49**	**55**
RMS 10.15 cents	**57**	**55**

It is immediately obvious that roughly a third of respondents listening to the control piece reported that piece to go out of tune. Roughly 50% of listeners to the test groups accurately reported each group to go out of tune.

Comparison of Test Groups with Control
Using a chi-squared test for independence, both test groups were compared to the control group. The null hypothesis in each case was that the presence of a drift in pitch would not cause a significantly different rate of reporting of drift by listeners.

CONTROL VS TEST PIECE 1

No Drift	30	61	91
RMS 4.76 cents	49	55	104
	79	116	195

The analysis gives a result of P=0.0447, therefore the null hypothesis is rejected – respondents *were* more likely to judge this test piece to have drifted in tuning.

While the proportion of people reporting a drift in this test piece had increased compared to the control piece, the absolute proportion of people who identified the change was close to half.

CONTROL VS TEST PIECE 2

No Drift	30	61	91
RMS 10.15 cents	57	55	112
	87	116	203

This analysis, again, gives a result of P=0.0103, rejecting the null hypothesis – respondents were also more likely to judge the test piece to have drifted in tuning than the control piece. The absolute proportion is still roughly half – similar to that of Test Piece 1.

Those reporting the test piece to stay in tune number 55 for both test pieces, allowing a convenient comparison at a glance. The number reporting the piece out of tune does rise slightly with a higher deviation in tuning, but is this rise statistically significant?

TEST PIECE 1 VS TEST PIECE 2

RMS 4.76 cents	49	55	104
RMS 10.15 cents	57	55	112
	106	110	216

Another treatment of both test groups returns a P value of 0.579, suggesting that no, the change in responses between the two test pieces isn't unlikely to be down to chance alone.

D. DISCUSSION

A third of the control group reported the control piece (which didn't experience any drift) to go out of tune. This could have been influenced by one of three things: 1) artifacts caused by the software used to retune the music; 2) misunderstanding of the term 'out of tune' possibly aggravated by the slight use of chromatic passing notes in the music; 3) respondents' suggestibility when listening for a specific effect.

The method of retuning the musical pieces was with Melodyne Editor pitch manipulation software. While each piece (even the control piece) had been subject to this audio treatment, any perceptible artifacts from it would have been stronger with higher deviation of the original note pitches. Due to this, it's entirely possible that these artifacts could have influenced listeners' responses in the same way as the actual tuning drifts themselves. A more appropriate way to achieve pitch control when using sampled instruments in this way would be to use MIDI pitchbend code throughout rather than ad hoc retuning of the audio.

A better way to further investigate this would be the uptake of a forced choice staircase experiment, in which different stimuli are presented, the tuning of which is based on the correctness of the respondent's previous answer, until a reliable, constant threshold for 50% success is reached. The results of the above study would predict this 50% point to lie between RMS 4 and 10 cents over the same timeframe.

E. REFERENCES

1. **Sethares, W A.** *Tuning, Timbre, Spectrum, Scale.* s.l. : Springer, 1998
2. **Isacoff, S.** *Temperament.* s.l. : Faber, 2002
3. **Duffin.** *How Equal Temperament Ruined Harmony.* s.l. : Norton, 2007
4. Yahoo Tuning List. [Online] [Cited: 07 27, 2013.] http://groups.yahoo.com/group/tuning/
5. Xenharmonic Wiki. [Online] [Cited: 07 27, 2013.] http://xenharmonic.wikispaces.com/
6. Facebook: Xenharmonic Alliance II. [Online] [Cited: 07 27, 2013.] https://www.facebook.com/groups/xenharmonic2
7. Redundant 12-tone Tunings. *21harmony.com.* [Online] [Cited: 07 28, 2013] http://www.21harmony.com/1/post/2013/04/redundant-12-tone-tunings.html
8. **Barbour, M J.** *Tuning and Temperament, a historical survey.* s.l. : Dover, 2004 (1951)
9. **H Fastl, E Zwicker.** *Psychoacoustics: Facts and Models.* s.l. : Springer-Verlag, 1999
10. *Frequency and intensity difference limens for harmonics within complex tones.* **Moore, Brian C. J.** 2, s.l. : J Acous. Soc. Am, 1984, Vol. 75
11. *Frequency discrimination as a function of frequency, measured in several ways.* **Moore, Brian C. J.** 4, s.l. : J Acous. Soc. Am., 1995, Vol. 97
12. 21harmony. *The Trouble with Melodyne Editor.* [Online] [Cited: 08 03, 2013] http://www.21harmony.com/1/post/2012/05/the-trouble-with-melodyne-editor.html