



<http://www.diva-portal.org>

This is the published version of a paper presented at *Nordic Acoustical Meeting*.

Citation for the original published paper:

Ternström, S., Nordmark, J. (1996)

Intonation preferences for major thirds with non-beating ensemble sounds

In: *Proc. of Nordic Acoustical Meeting: NAM'96*, F2 (pp. 359-365). Helsinki

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-227976>



INTONATION PREFERENCES FOR MAJOR THIRDS WITH NON-BEATING ENSEMBLE SOUNDS

Sten Ternström, Jan Nordmark

Royal Institute of Technology, Dept. of Speech, Music and Hearing
Box 70014, S-100 44 Stockholm, Sweden
sten@speech.kth.se

1 INTRODUCTION

The most important distinguishing characteristic of the Pythagorean, the pure, and the equal-temperament scales is the size of the major and minor thirds. While the octave is identical in these tuning systems, and the fourth and fifth differ only by small amounts, the thirds vary considerably. The only intervals considered to be consonant in the Pythagorean system were the octave, the fifth, and the fourth, which could all be derived by the simple operation of the numbers 1 to 4. The divisions were based on a philosophical or numerological principle rather than on the acceptability of the sounds of the intervals themselves. As a consequence of this principle the frequency ratio of the interval of the major third was the rather complex one of 81:64.

The increasing use of the third in final cadences in the 16th century became difficult to reconcile with its status as a dissonance in the Pythagorean system. The solution to this problem came with Zarlino's extension of the basic numbers for interval division to include the numbers 5 and 6, thus permitting the major third to be defined as the integer ratio 5:4, and the minor third as the ratio 6:5. Zarlino's system permitted an unequivocal rank ordering of consonances among the intervals: the simpler the numerical relationships, the greater the consonance, and intervals with ratios including the number 7 were definitely dissonant. In the Pythagorean system the ratio of the third was actually more complex than that of the major second, although there could be no doubt that the third was the more consonant.

It is therefore hardly surprising that the system of pure tuning came to be considered theoretically superior to the Pythagorean one. This view was reinforced by the discovery of the overtone series in the complex tone produced by musical instruments. The series contains overtones in the ratio 5:4, and a musical chord consisting of a pure third and a pure fifth was considered to be the "chord of nature." Theories of harmony from Rameau to Schenker are mostly based on the presumed naturalness of the major chord, and implicitly on the system of pure tuning. Helmholtz' linking of dissonance to the presence of beats also contributed strongly to the enduring belief in the superiority of pure tuning, as beats were minimized in intervals tuned in a simple relationship. Pure tuning had problems of its own, however, and could not be used for instruments with fixed tuning without causing unacceptable intervals. The invention of equal temperament overcame this problem, but at the cost of intervals which

were all, except for the octave, somewhat mistuned. The difference was particularly large for the major and minor thirds. Equal temperament was therefore considered to be at best a compromise, and in the case of the thirds, a particularly objectionable one.

There is, however, scant evidence in support of the idea of the naturalness of pure tuning. Measurements of the actual intonation preferences in solo and ensemble performances instead point to equal temperament and Pythagorean tuning rather than pure tuning (for a recent review, see Loosen, 1995 [1]). Particularly relevant is the study by Lottermoser and Meyer [2] on intonation in choral singing as measured from recordings. The advantage of such measurements is the absence in choral singing of any cues from beats that might influence the preferred size of the intervals. Beats do of course occur in choral sounds, but they are too abundant and irregular to act as intonation cues. In writings on intonation the idea is frequently advanced that choirs should or do use their freedom of intonation to achieve pure tuning [3, 4]. The results of Lottermoser and Meyer do not support this idea. They found the major thirds consistently to be even larger than in equal temperament. Similarly, Hagerman and Sundberg [5] studied intonation in barbershop quartets, and found large major thirds, even though there was only one voice per part.

For the present investigation, musically expert subjects with choral and/or orchestral experience were asked to adjust simultaneous major thirds to their preference. The sounds were synthesized ensemble sounds that did not give rise to regular beats.

2 METHOD

2.1 Stimulus sounds

The stimulus sounds were designed to resemble an ensemble of six musical instruments of a neutral character, somewhat reminiscent of violas. Of the six voices, three produced the prime note (P_1 , P_2 , P_3) and three produced the major third above it (T_1 , T_2 , T_3). A pseudo-stereophonic presentation was achieved by feeding voices P_1 , P_2 , T_1 and T_2 to the left channel and voices P_3 , T_3 and T_2 to the right channel. Voices P_2 and T_2 were given half amplitude in both channels. This resulted in a convincing ensemble sound that spread across the left-right baseline.

Each instrument voice was produced by passing a sine wave at the fundamental frequency through a peak follower with a release time constant of 4 ms. Filters were used to achieve a timbre that was not too boomy nor too sharp: high-pass at 630 Hz, low-pass at 1560 Hz, both 6 dB per octave (figure 1).

The fundamental frequency F_0 of each voice was made unsteady by using a model of human voice flutter [7]. In this model, the nominal F_0 is perturbed with a flutter signal: a white noise that has passed through a second-order resonant filter with a resonance frequency of 4.3 Hz and a bandwidth of 4.0 Hz. Each voice had its own flutter noise generator, independently seeded (figure 2).

The amplitude of the flutter is important. If it is too small, regular beats can still be heard, while if it is too large, the sensation of a definite pitch may suffer. The flutter amplitude was therefore chosen in the following way. The flutter amplitude was first set to zero, thus yielding straight tones and a composite sound that was entirely devoid of an ensemble effect. The nominal F_0 of the P voices was set to 220 Hz. The nominal F_0 of the T voices was then

adjusted so as to generate strong beats (390 cents above, which is slightly larger than the pure major third at 386 cents). The amplitude of the flutter was then increased gradually until regular beats were no longer discernible. This resulted in a long-term standard deviation in F_0 of 9 cents, called the *flutter level*. Typical flutter levels in vowels sustained by choir singers are 10-15 cents [8]. The flutter level was kept at 9 cents for all six voices and for all dyads.

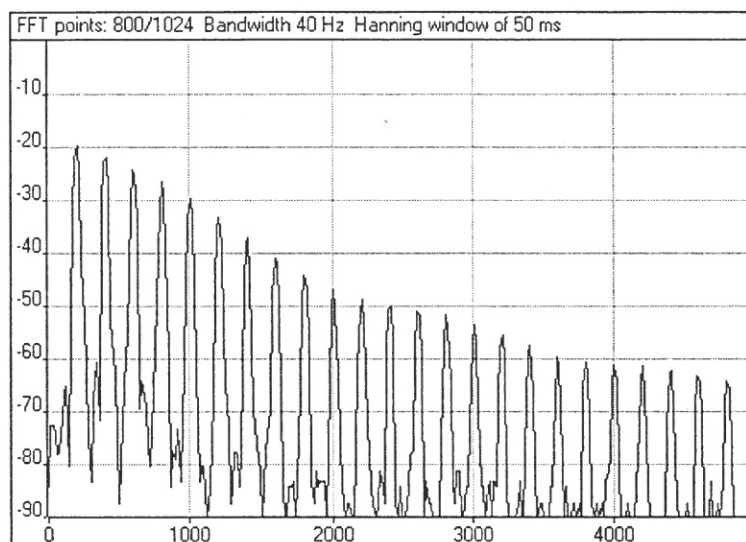


Figure 1: A spectrum section of one 'voice' in isolation, without flutter. The spectrum of the sum of voices would fluctuate rapidly due to the quasi-random beating of all partials.

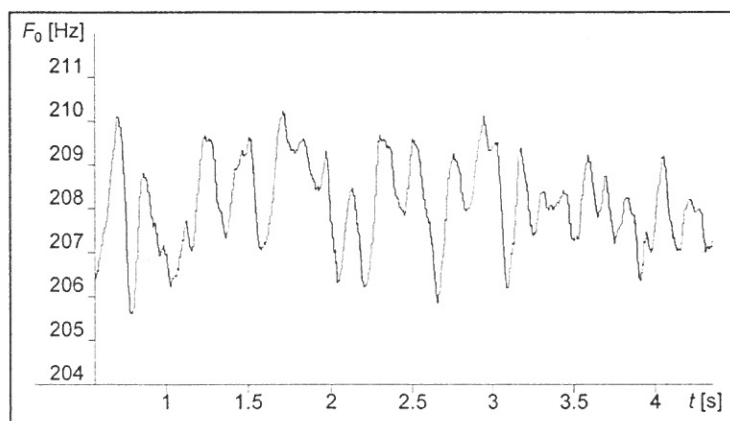


Figure 2: F_0 extraction of one voice at 208 Hz, showing the effect of flutter at large magnification.

2.2 Stimulus presentation and procedure

The stimulus tones were presented over headphones (Sennheiser HD414). The tones were gated 3 seconds at a time, with a pause of 1.5 seconds and a rise and fall envelope as produced by a first-order low-pass filter at 2 Hz.

The subjects could adjust continuously the nominal F_0 of the T voices over a range of 400 ± 150 cents relative to the P voices. This was done using a free-standing multi-turn rotary control knob, with a resolution of better than 0.1 cents. This control knob (and thus the nominal F_0 initially presented to the subjects) was given a random numerical bias, such that the knob setting chosen for one dyad would seldom be appropriate for the next dyad. The subject would adjust the size of the major third interval to his or her satisfaction, and then press a key. This would cause the control program to register the chosen interval size, and then advance to the next dyad. However, if the chosen major third interval were smaller than 350 cents or larger than 450 cents, the program would not advance to the next dyad, thus indicating an error to the subject. Preliminary trials had showed that this was necessary – a few pilot subjects occasionally slipped into the minor third.

The major third dyads were presented at ten different pitches with the F_0 of the prime ranging pseudo-randomly in semitone steps from 155 Hz to 392 Hz. Each dyad was replicated, making a total of 20 dyads. The sequence of dyads was the same for all subjects. There was no time limit for the adjustment. The experiment was conducted in a quiet office environment. There were three suitably equipped PC workstations, so that three subjects at a time could be active.

2.3 Subjects

Sixteen subjects participated, selected for their level of musical training. Eleven of them were undergraduate students of choral pedagogy or choral conducting, and five were orchestra musicians.

2.4 Technical implementation

The entire sound synthesis was implemented in a real-time model (figure 3), built using the system *Aladdin Interactive DSP* 1.0 from AB Nyvalla DSP, Stockholm [9]. The model ran on a Texas Instruments 32-bit floating point digital signal processor mounted on a commercial PC add-in board from Loughborough Sound Images, UK. The model and the experimental procedure were controlled automatically by a custom experiment manager program written in Microsoft Excel 5.0. The control knob was a that of a multiturn potentiometer acting as a simple voltage divider for a small battery. Its output voltage controlled the deviation from equal temperament via an A/D-converter. The synthesized stereo sound was output via two high-quality D/A converters on the board.

3 RESULTS

Each subject typically took 10-15 minutes to complete the 20 dyads. **The average preferred size of the major thirds was 395.4 cents**, with an overall standard deviation of 7.3 cents. This is closer to equal temperament (400) than to pure intonation (386). Figure 4 shows the preferred size rank ordered by subject. There was no systematic effect of F_0 on the preferred size of the major third.

A consistency measure was defined for each subject, as the mean of the absolute value of the difference between the replicates of the same dyad. This value would be small if the subject tended to choose the same intonation on both occasions. Figure 5 shows the same data as figure 4, but subjects are rank ordered by consistency. It can be seen that even the most consistent subjects had quite different preferences.

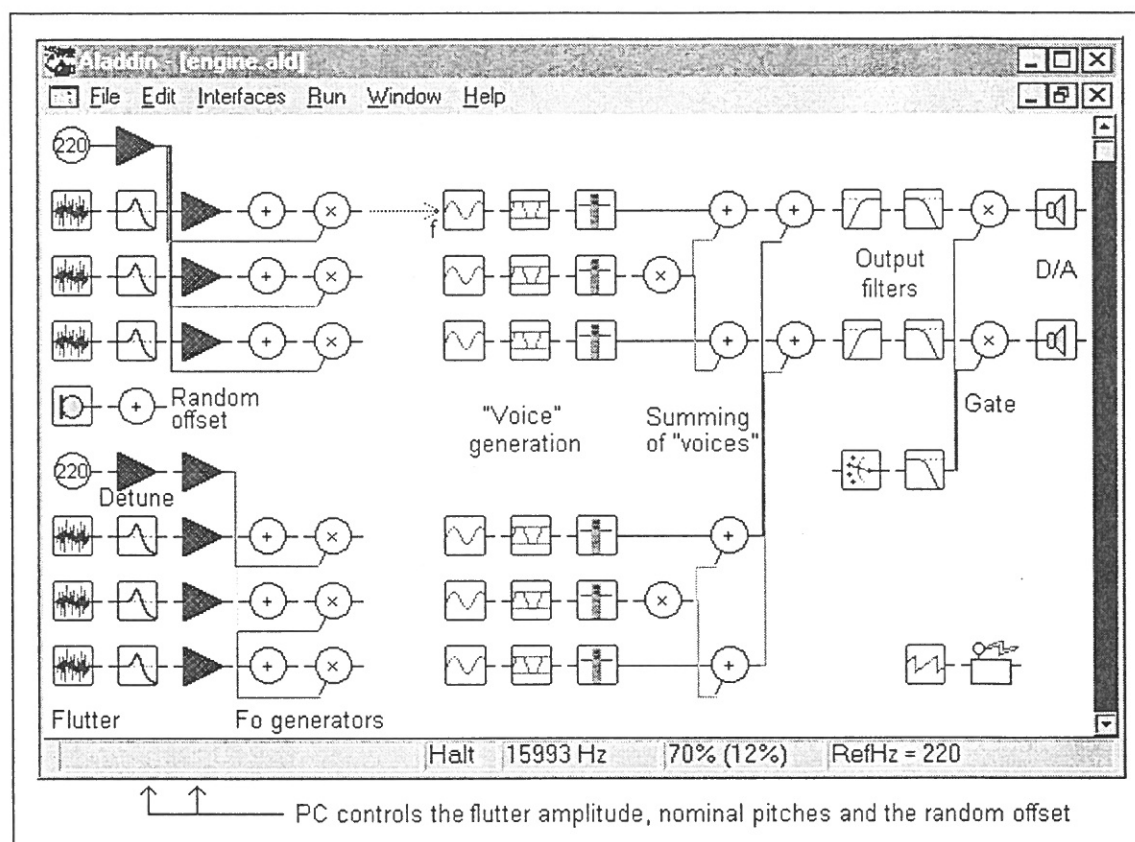


Figure 3: The signal generation model used to produce the ensemble sounds. The model ran in real time on a digital signal processor which was controlled by a spreadsheet program on the host PC.

4 DISCUSSION

The results show no sign of preference for pure intonation (386 cents). Rather, there is a wide spread of subject preferences ranging from 388 cents to 407 cents, and the response distributions for some subjects do not even overlap. Subject 4 (405 cents) remarked that his tuning strategy was first to tune the major third much too sharp and then to decrease it until he found it acceptable. This strategy may account for his high intonation. Subject 13 (407 cents) remarked that as a choir singer she had a habit of 'pitching up' because some of her section colleagues were prone to sing flat (in her opinion).

The present experiment shows that 16 musically experienced subjects had quite different subjective preferences for the size of major third dyads; and that, in the absence of regular beats, so-called pure intonation (386 cents) does not seem to be universally desirable.

ACKNOWLEDGEMENTS

This work was supported by the Swedish Natural Science Research Council (NFR). Gustav Levander kindly assisted us in running the experiment. We are grateful to Eric Prame for valuable discussions.

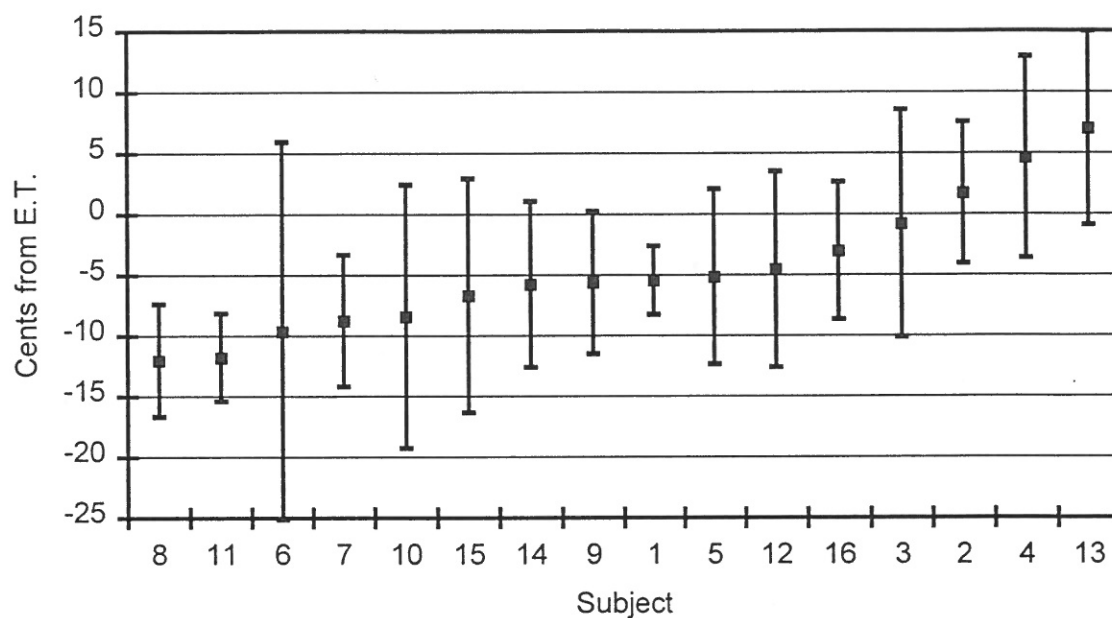


Figure 4: Subjects' preferred intonations of major thirds, represented as cents deviation from 400 cents (equal temperament). Each point is a mean of twenty dyads, with vertical bars representing the standard deviation for that subject.

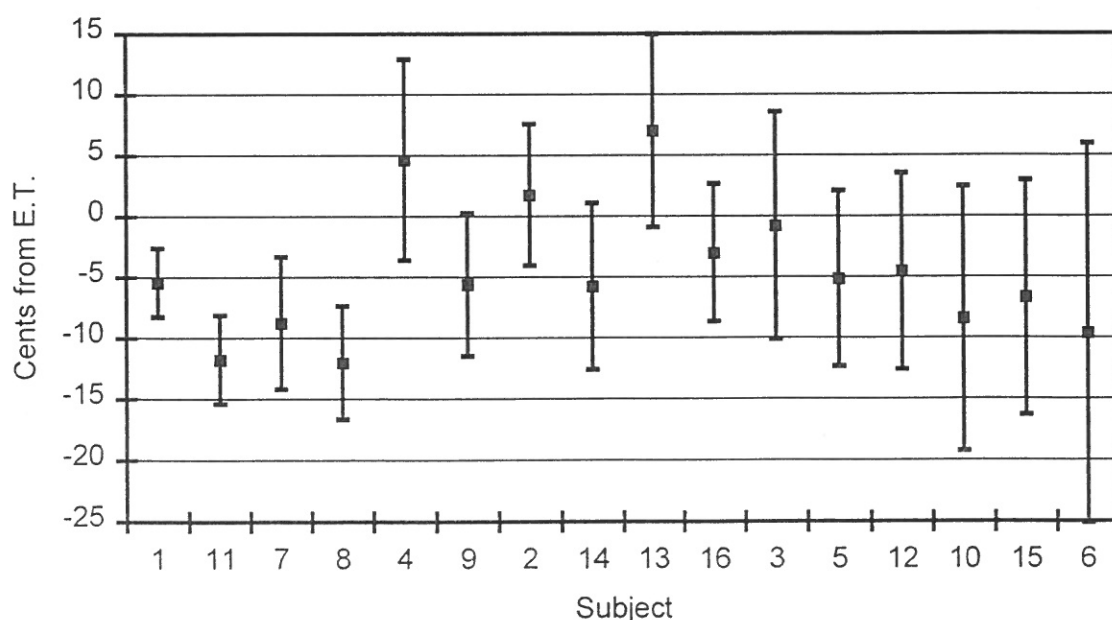


Figure 5: The same data as in figure 4, but subjects are rank ordered by consistency (best to worst).

REFERENCES

1. Loosen F. 1995. The Effect of Musical Experience on the Conception of Accurate Tuning. *Music Perception*, vol. 12, no. 3, pp. 291-306.
2. Lottermoser W, Meyer Fr-J. 1960. Frequenzmessungen an gesungenen Akkorden. *Acustica*, vol. 10, pp. 181-184.
3. Benade A H. 1976. *Fundamentals of Musical Acoustics*, p. 295. Oxford University Press.
4. Pickering J C. 1995. *Acoustically pure intonation in a cappella vocal music*. Thesis for Master of Music, Australian National University, February 1995.
5. Hagerman B, Sundberg J. 1980. Fundamental frequency adjustment in barbershop singing. *J Res Singing*, vol 4., no. 1, pp 3-17.
6. Ternström S, Sundberg J. 1988. Intonation precision of choir singers. *J Acoust Soc Am*, vol. 84, no. 1, pp. 59-69.
7. Ternström S, Friberg A. 1989. Analysis and simulation of small variations in the fundamental frequency of sustained vowels. *STL-QPSR* 3/89, 1-14.
8. Ternström S. 1993. Perceptual Evaluations of Voice Scatter in Unison Choir Sounds. *J Voice*, vol. 7, no. 2, pp. 129-135.
9. AB Nyvalla DSP. 1995. *Aladdin Interactive DSP*. Available from AB Nyvalla DSP, Roslagsvägen 101, bldg 15, S-104 05 Stockholm, Sweden. Internet: info@nyvalla-dsp.se. This software system was written by author S.T. and Lennart Neovius.