A COMPUTATIONAL COMPARISON OF THEORY AND PRACTICE OF SCALE INTONATION IN BYZANTINE CHANT

Maria Panteli[1] **& Hendrik Purwins**[2,3]

[1]Department of Computer Science, University of Cyprus, Cyprus
[2]Neurotechnology Group, EE & CS, Berlin Institute of Technology, Germany
[3]Sound and Music Computing Group, Aalborg University Copenhagen, Denmark
{m.x.panteli, hpurwins}@gmail.com

ABSTRACT

Byzantine Chant performance practice is quantitatively compared to the Chrysanthine theory. The intonation of scale degrees is quantified, based on pitch class profiles. An analysis procedure is introduced that consists of the following steps: 1) Pitch class histograms are calculated via non-parametric kernel smoothing. 2) Histogram peaks are detected. 3) Phrase ending analysis aids the finding of the tonic to align histogram peaks. 4) The theoretical scale degrees are mapped to the practical ones. 5) A schema of statistical tests detects significant deviations of theoretical scale tuning from the estimated ones in performance practice. The analysis of 94 echoi shows a tendency of the singer to level theoretic particularities of the echos that stand out of the general norm in the octoechos: theoretically extremely large scale steps are diminished in performance.

1. THE OCTOECHOS AND THE CHRYSANTHINE THEORY

Byzantine Chant is the Christian liturgical song of the Eastern Roman Empire (*Byzantium*) that gradually emerged from the Roman Empire from the 4th century on. Byzantine Chant has been the dominant liturgy of the Eastern orthodox Christianity. Referring to various theoretic accounts on Byzantine Chant, Zannos in [18] argues that *'none of them can be said to correspond with contemporary empirical study'*.

The main analysis tool used was a pitch class profile [17] with high bin resolution, extracted from audio recordings with the aid of specifically designed algorithms. These were applied on a music collection of 94 Byzantine Chants labeled after the scale they are performed in. The overall behavior and consistency of empirical scale degree tuning was computed and contrasted to theory through a series of tests and experiments.

According to Mavroeidis [12] and Thoukididi [16], a mode (singular: *echos*, plural: *echoi*) in Byzantine Chant is defined by the following five characteristics: 1) the *scale degree steps (SD)* between consecutive scale degrees, 2) the most prominent scale degrees (two or three scale degrees out of which I - III and I - IV scale degrees are the most reoccurring pairs), 3) a short introductory phrase that marks the reference tone, 4) the cadences in the middle and the end of a phrase, and 5) the modulations (alterations) applied to particular scale notes depending on whether they are reached by an ascending or a descending melody. Subject to a reform in the 1880s in particular concerning the sizes of the intervals, the Chrysanthine notation method (from the 1820s) is used in the official chant books of the Greek Orthodox Church up to now [11]. In Chrysanthine theory, the octave is divided into 72 equal partitions, each of 16.67 cents (singular: *morio*, plural: *moria*). The scale degree steps are measured in multiples of a morio (cf. Table 1). According to this theory there are in total eight basic echoi, a system also referred to as *octoechos* ('eight' + 'mode'). These eight modes occur in pairs of *authentic* and corresponding *plagal* modes: *First Authentic, Second Authentic, Third Authentic, Fourth Authentic, First Plagal, Second Plagal, Grave, Fourth Plagal*. The plagal mode has a different reference tone (tonic) than its authentic counterpart, usually a perfect fifth lower than the one of the authentic mode [1] but it may share the same scale step sequence. Furthermore, both differ in melodic characteristics. The scale degree steps may vary according to the chant genre (*Heirmoi, Stichera, Papadika*) [12, 16]. Our study is limited to the basic and simplest echos scales [2]. We will not consider the fact that scale degree steps of an echos can be modulated (altered) based on the melodic characteristics of a chant or other criteria [12, 16] (cf. Section 4).

2. MUSIC CORPUS

The music corpus analysed in this study consists of recorded monophonic songs from the album series of Protop-

---

[1] The *Grave* shares the same scale degree steps with the Third Authentic as well as the same reference tone.

[2] Chants of Heirmoi genre of Second, Heirmoi of Second Plagal, Papadika of Grave, Stichera and Papadika of Fourth, and Papadika of Fourth Plagal use different scale steps than the eight basic echoi thus omitted in this study (cf. Table 1).

Echos	Chant Type	Step (in Moria) between Scale Degrees							
		I	II	III	IV	V	VI	VII	I
First	All chants	10	8	12	12	10	8	12	
First Plagal	All chants	10	8	12	12	10	8	12	
Second	Stichera/Papadika	8	14	8	12	8	14	8	
Second Plagal	Stichera/Papadika	6	20	4	12	6	20	4	
Third	All chants	12	12	6	12	12	12	6	
Grave	Heirmoi/Stichera	12	12	6	12	12	12	6	
Fourth	Heirmoi	8	12	12	10	8	12	10	
Fourth Plagal	Heirmoi/Stichera	12	10	8	12	12	10	8	

Table 1. The scale structure of the eight modes (octoechos) measured in multiples of a morio. The two tetrachords are indicated.

saltes Georgios Kakoulides [8], Protopsaltes Ioannis Damarlakis [3], Protopsaltes Panteleimon Kartsonas [9], and Protopsaltes Dimitrios Ioannidis [7]. From the total of 94 recordings, 13 are in the First Authentic echos, 15 in First Plagal, 6 in Second Authentic, 6 in Second Plagal, 18 in Third Authentic, 9 in Grave, 10 in Fourth Authentic and 17 in Fourth Plagal.

3. COMPUTATIONAL ANALYSIS PROCESS

To study scale degree pitch empirically, pitch modulo octave histograms are investigated. Built on pitch histograms, pitch class profiles have been applied to detect key and tone centres in classical Western music [6, 14]. The authors in [2] adapted the latter approach to Raag recognition. Bozkurt [1] proposed a method to extract the tuning of a scale, applied to Turkish maqam. Moelants et al. [13] introduced a peak picking heuristics to extract the scale tuning from a modulo octave pitch histogram of African scales. Serrà et al. [15] used pitch class histograms to study scale tuning in Hindustani and Carnatic music, investigating whether this music follows equal temperament rather than just intonation.

The procedure followed in this article is summarized in Figure 1. First, the pitch ($f0$) trajectory is extracted from each audio recording. A pitch histogram is computed, compressed into one octave and smoothed. Peaks are extracted from the histogram and are employed in tonic detection. The pitch trajectory is then aligned to the estimated tonic. For recordings of a particular echos, the aligned pitch trajectories are used to compute the *echos histogram*. Peaks are then detected in the echos histogram, and peak locations are mapped to theoretical scale degree pitches. Pitches around the selected peak locations are used to determine a neighbourhoud of pitches around the empirical scale degrees. Finally, a sequence of statistical tests is used to compare the estimated practical scale tuning with the theoretical ones.

3.1 Pitch Trajectory via F0 Detection

In the current study we use the f_0 estimation Yin algorithm [4]. Considering the melodic characteristics of the analysed music as well as the particularities of the singing voice, the following post processing filters were designed:

noise, silent gaps, octave/fifth error. As a result, a trajectory of N_m estimated pitches $p = (p_1, \ldots, p_{N_m})$ is generated for each recording m. Three experts evaluated the estimated pitch trajectory of 20 excerpts, presented as a sythesized version of the original melody, with an average 4.2 of 5(max).

3.2 Pitch Histogram via Kernel Smoothing

For the vector p, we define the *pitch histogram*

$$c_k^{p,b} = \Sigma_{n=1}^{N} q_r\left(\frac{p_n - b_k}{h}\right) \qquad (1)$$

with the *rectangular kernel function*

$$q_r(u) = \begin{cases} 1 & \text{if } |u| \leq \frac{1}{2} \\ 0 & \text{otherwise} \end{cases} \qquad (2)$$

for *K center bins* $b = (b_1, \ldots, b_K)$, being multiples $b_k = (k-1).h + p^*$ of bin widths h and an offset p^*. The Chrysanthine theory divides the octave into 72 equal partitions and we multiply this division by 3 arriving at $K = 216$ center bins in the range of one octave, thereby yielding sufficient bin resolution and robustness. The choice of h is critical for the subsequent stage of peak picking, since a too high h can eliminate relevant peaks whereas a too small h can create spurious peaks in the pitch histogram.

Although a large h increases the smoothness of the pitch histogram, discontinuities in the histogram remain. These discontinuities are artefacts due to the partitioning of the pitches in a discrete set of predefined bins. The sharp-edged rectangular kernel function in (1) is replaced by a smooth *Gaussian kernel* function yielding the equation

$$c_k^{p,b} = \Sigma_{n=1}^{N} \frac{1}{\sqrt{2\pi h^2}} e^{\frac{-\|p_n - b_k\|^2}{2h^2}}. \qquad (3)$$

The selection of the appropriate smoothing parameter h is guided by the task the histogram is used for. For the estimation of the SD tuning a relatively high smoothing factor is employed to avoid spurious peaks in a too detailed histogram. To determine an adequate h, the following assumption is made: Byzantine theory defines 4 moria (67 cents) as the smallest SD interval. Choosing the quartertone ($\delta_{\min} = 50$ cents) as the smallest acceptable distance between two histogram peak locations allows for a margin
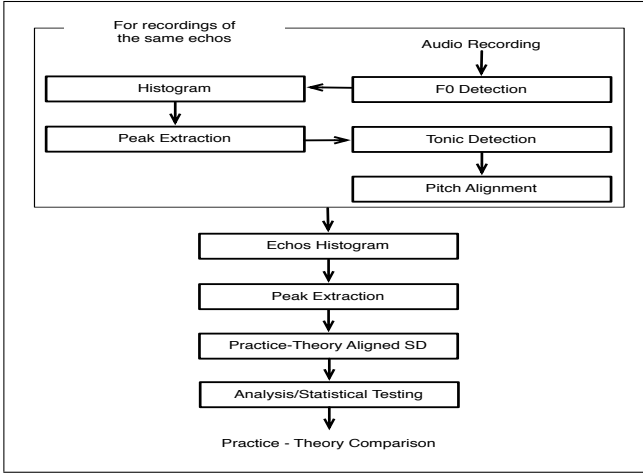
Figure 1. Flow diagram of analysis process of a recording yielding a theory-practice comparison of the scale tuning (SD=Scale Degree).

for investigating the deviations between theory and practice. Experimenting with h, this assumption is satisfied when h is set to 18 cents.

3.3 Peak Extraction

Using a peak extraction algorithm, from the smoothed pitch class histogram $c^{p,b}$, U peaks (λ, π) can be detected consisting of *peak locations* $\lambda = (\lambda_1, \ldots, \lambda_U)$ and *peak amplitudes* $\pi = (\pi_1, \ldots, \pi_U)$. In [13] the authors propose a number of heuristics to be used in peak picking, such as the size and height of peaks and intervals between peaks. Our proposed algorithm iteratively chooses the peak position $\lambda_u = b_{k_u}$ with maximum peak height $\pi_u = c_{k_u}$, then removing the potential location candidates in a $\pm\delta_{\min}$ neighborhood around the selected peak position, to choose the next peak until U peaks are picked. U is defined as follows: According to Byzantine theory each echos has at least 7 scale notes. In addition, Byzantine theory knows of note alterations. To account for further intonation variants in practice, we set $U = 12$. $\delta_{\min} = 50$ cents is defined as the minimum neighborhood.

3.4 Tonic Detection

To make the first bin b_1 corresponding to the tonic, the histogram has to be circularly (modulo K) shifted by p_0, the pitch of the tonic. The authors in [5] calculated the cross-correlation between all scale degree prototypes and all circularly shifted versions of a smoothed histogram. The pitch shift p_0 that gives the maximum cross-correlation is the estimated tonic and is used to circularly shift the second histogram. According to theory, the tonic of the Byzantine echos considered in this study is stated at the end of the phrase. Our tonic detection algorithm computes the pitch of the last phrase note from the onset and frequency information assuming that 1) the last note in the recording is the last note of the melodic phrase and that 2) the final phrase note lasts for at least half a second.

To compensate with inaccuracies [5], the tonic detection algorithm integrates pitch information from a set of maximum three onsets detected at the end of the phrase. Conditions apply to decide which of the three onsets correspond to the last phrase note considering vibrato and ornamentation (the main incaccuracies in this case), and the pitch is then estimated as an average of these. As a final step, the estimated pitch of the last phrase note is refined to the closest histogram peak that represents the closest empirical scale degree. Three experts found a wrongly automatically estimated tonic in 2 out of 20 excerpts.

3.5 Practice - Theory Comparison

For recordings m, $1 \leq m \leq M_i$, of echos i, the pitch trajectories $p^m = (p_1, \ldots, p_{N_m})$ are aligned to the estimated tonic and the *echos histogram* $C_i = c_k^{p^{1 \leq m \leq M_i}, b}$ is computed. The peak extraction algorithm applied on C_i yields the peak locations $\lambda = (\lambda_1, \ldots, \lambda_U)$ and peak amplitudes $\pi = (\pi_1, \ldots, \pi_U)$ of echos i.

3.5.1 Practice - Theory Aligned Scale Degrees

The normalized theoretical scale degrees are defined as a set of scale degree pitches (locations) $\nu^\theta = (\nu_1^\theta, \ldots, \nu_L^\theta)$, with normalization $\nu_0^\theta = 0$ (in cents) and $L = 7^{\,3}$. From the preselected peaks (λ, π) of echos i, with pitches $\lambda = (\lambda_1, \ldots, \lambda_U)$ and amplitudes $\pi = (\pi_1, \ldots, \pi_U)$ we estimate the practical scale degree pitches $\nu^\pi = (\nu_1^\pi, \ldots, \nu_L^\pi)$. First, only those peaks (λ, π) are selected that correspond to theoretic scale notes ν^θ, i.e. those pitches λ_u that lie closer to and within a d distance, from a ν_l^θ, yielding the pitch vector ν. The distance d is set to 150 cents ($\frac{3}{4}$ of a whole tone) to allow enough margin of deviation between empirical peaks and theoretical scale notes. If two or more pitches λ_u fulfill this condition, the peak with highest amplitude π_l is selected, since it corresponds to a more prominent note such as a scale note. The estimated scale degree pitch $\hat{\nu} = (\lambda_{t_1}, \ldots, \lambda_{t_{L'}})$ is defined by $t_l = \arg_{|\lambda_t - \nu_l| \leq d} \max(\pi_t)$. If no pitches λ_u lie within a d range around the theoretical pitch ν_l^θ, t_l is not defined and the estimated scale $\hat{\nu}$ has less degrees than the theoretical scale ν^θ, $(1 \leq L' \leq L)$.

From the estimated scale $\hat{\nu}$ and for each recording m, $1 \leq m \leq M_i$, of echos i, pitches $\hat{p_l}^m$ are selected from the trajectory $p = (p_1, \ldots, p_{N_m})$ that lie within 2-moria distance of the estimated scale degree pitches $\hat{\nu}_l$, $1 \leq l \leq L'$. The 2-moria distance defines half the size of the smallest theoretical scale interval (cf. Table 1). Statistical testing is applied to check whether pitches $\hat{p_l}^{1 \leq m \leq M_i}$ are derived from distributions with mean equal to the theoretical scale degree pitches ν_l^θ for $l = 1, \ldots, L'$. For the pitches $\hat{p_l}^m$ and theoretical scale degree ν_l^θ, the *Mean Deviation*, $mean(\hat{p_l}^m - \nu_l^\theta)$ is also computed.

3.5.2 Statistical Testing

To assess the deviation between theoretic and practical scale degree pitches and steps, we apply a chain of tests as an

---

[3] Exceptions with $L = 8$ theoretical scale degree pitches exist in some variations of echoi not included here.

analytical instrument. As a first step, the Shapiro-Wilk test is applied, to determine whether the scale degree pitches $\hat{p}_l^{1 \le m \le M_i}$ are normally distributed across all M_i instances of the same echos i. If the p value is above significance level $\alpha_n = 0.05$, we assume normal distribution and apply the t-test to $\hat{p}_l^{1 \le i \le I}$. The t-test hypothesis is formulated that for echos i, the l-th estimated scale degree pitches $\hat{p}_l^{1 \le m \le M_i}$ are derived from a distribution with mean equal to the l-th theoretical scale interval ν_l^θ. In case the Shapiro-Wilk rejects the normality hypothesis the Wilcoxon Signed-Rank test is applied instead. The significance level α_i for an individual test is set based on the *Bonferroni correction*; since $n = 48$ individual tests are applied and $\alpha = 0.05$ defines the confidence interval of the whole family of tests, each hypothesis is tested at the significance level $\alpha_i = \frac{\alpha}{n} = \frac{0.05}{48} \approx 0.001$.

If $p < \alpha_i$, for the probability p of observing $t_{i,l}$ under the null hypothesis, we reject the null hypothesis and conclude that theoretical scale degree pitch deviates significantly from the practical scale degree pitch. In addition, the histogram around an empirical scale degree pitch could be characterised by parameters such as variance, skewness, kurtosis, following [10].

4. RESULTS

The pitch histograms of all recordings of all echoi can be found in Figure 2. The tuning of the Byzantine scales has been investigated by comparing the pitches of empirical and theoretical scale notes. A series of statistical test was employed to determine for which scale notes practice and theory of tuning deviate. For all recordings of a given echos, the pitches $\hat{p}$ around the estimated scale degrees $\hat{\nu}$ (at the peaks of the histogram) are gathered. The Shapiro Wilk test on normality performed for pitches $\hat{p}$ showed that the normality hypothesis is rejected for all estimated scale degree pitches. The Wilcoxon Signed-Rank Test (W-test) with significance level $\alpha_i = 0.001\%$ (based on the Bonferroni correction) is therefore applied to test the null hypothesis that the median of the empirical pitches is the same as the theoretical pitch of a particular scale degree in a particular echos.

Table 2 reveals that the majority of scale degrees of all echoi differ significantly from theory. The null hypothesis was not rejected for the scale degrees V of First, VII of First Plagal, VI of Second Plagal, and II of Third and Grave. The Second, Fourth and Fourth Plagal echoi have all their empirical scale degrees significantly deviating from theory. The VI. scale degree of the First and First Plagal has a relatively large mean deviation value with negative sign, i.e., the empirical scale degree pitch is smaller than the theoretical one. According to theory, alterations of these echoi apply that diminish particularly the VI. scale degree when the melody is descending. Other large mean deviations appear for the VII. scale degree of Second Authentic as well as the III. and VII. scale degrees of Second Plagal. These scale degrees are reached by relatively large theoretical scale intervals; the VII scale degree of Second Authentic is reached by the VI-VII scale step of 14 moria

whereas the III. and VII. scale degrees of Second Plagal are both reached by a scale step of 20 moria (cf. Table 1). The empirical scale degrees appear with negative deviation, i.e. large scale intervals are diminished in performance. Fourth is the only echos, in which the first tetrachord is extended by two moria comparised to the other echoi. In practice, the V. scale degree of Fourth Authentic is significantly diminished with respect to the theoretical scale degree pitch. An interpretation could be that the singer tends to diminish the abnormally high tetrachord pitch of this echos. In the same echos, also scale degree step VII-I (theoretically 10 moria) tends to be diminished towards the more common step VII-I of 6 moria.

One may argue that the assignment of empirical peaks to the nearest theoretical peak may introduce a dependency between the theoretical and practical peak positions and therefore bias the test. However this bias is limited due to the following reasons: An informal inspection reveals that the empirical peaks are relatively close to the theoretical peaks, and, in all but few exceptions, there are exactly seven empirical peaks that correspond to the theoretical scale degrees I-VII.

5. CONCLUSION

In this paper, a new method has been introduced to empirically study the tuning of scale degrees. The method has been used to investigate to what degree performance practice of Byzantine Chant follows the widely known Chrysanthine theory. The theoretic hypotheses have been tested on a corpus of recordings of chants of the octoechos. A combined method of pitch estimation with appropriate post filtering has been applied to the recordings. Among the novel methods proposed here are the histogram computation algorithm that comprises the use of Gaussian kernel and suitable tuning of the smoothing factor. The analysis gives support to the conjecture that the singer levels the extreme step and scale degrees particularities within the octoechos in performance practice of Byzantine Chant. The methodology introduced here has applications to a wide range of oral music traditions.

6. ACKNOWLEDGEMENTS

7. REFERENCES

[1] B. Bozkurt. An Automatic Pitch Analysis Method for Turkish Maqam Music. *Journal of New Music Research*, 37(1):1–13, March 2008.

[2] P. Chordia and A. Rae. Raag recognition using pitch-class and pitch-class dyad distributions. In *Proceedings of ISMIR*, pages 431–436, 2007.

Echos	Feature	Scale Degrees					
		II	III	IV	V	VI	VII
First (13)	p (W-test)	0	0	0	0.0032	0	0
	Mean Deviation	-16.56	11.25	10.45	0.84	-43.89	5.79
First Plagal (15)	p (W-test)	0	0	0.0001	0	0	0.8767
	Mean Deviation	28.06	-5.65	-0.91	4.98	-50.06	-0.06
Second (7)	p (W-test)	0	0	0	0	0	0
	Mean Deviation	5.79	-5.56	-6.55	10.36	20.95	-49.64
Second Plagal (7)	p (W-test)	0	0	0	0	0.1592	0
	Mean Deviation	11.20	-44.32	5.06	10.79	0.47	-138.47
Third (18)	p (W-test)	0.0048	0	0	0	0	–
	Mean Deviation	0.45	-21.70	-11.19	-5.48	-11.16	–
Grave (8)	p (W-test)	0.0233	0	0	0	0	–
	Mean Deviation	-0.44	-27.81	-11.05	-17.42	-10.53	–
Fourth (11)	p (W-test)	0	0	0	0	0	0
	Mean Deviation	5.27	5.58	-5.21	-39.14	11.13	38.85
Fourth Plagal (16)	p (W-test)	0	0	0	0	0	0
	Mean Deviation	22.06	11.17	16.63	17.86	17.00	22.25

Table 2. Significance of scale degree pitch deviation between practice and theory for all echoi (number of instances in brakets). Since all echoi are aligned to pitch 0 for scale degree I, only II-VII are shown. For scale degrees II-VII the p-value of the test statistic and the mean pitch deviation (in cent) between practice and theory are indicated. The p-value of the Wilcoxon Signed-Rank test (W-test) is denoted. Zero p-values correspond to values smaller than 10^{-4}. Practice-theory deviations (cf. Table 1) greater than two moria are colored and discussed in the text.

[3] I. Damarlaki. *Pws tha mathw na psallw praktika [Learn how to perform the Byzantine chants] [Audio CD]*. Polychronakis, Crete, 1999.

[4] A. de Cheveigne and H. Kawahara. YIN, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America*, 111(4):1917, 2002.

[5] A. C. Gedik and B. Bozkurt. Evaluation of the makam scale theory of arel for music information retrieval on traditional turkish art music. *Journal of New Music Research*, 38(2):103–116, 2009.

[6] E. Gómez. *Tonal Description of Music Audio Signals*. PhD thesis, Universitat Pompeu Fabra, 2006.

[7] D. Ioannidis. *H theoria tis Byzantinis mousikis stin praxi [Theory of Byzantine music in practice]*. Ioannidis Dimitrios, Athens, 2005.

[8] G. I. Kakoulidis. *Pws na vriskoume praktika tous okto ihous tis Byzantinis mousikis [How to estimate empirically the eight modes of byzantine music]*. Athens, 1999.

[9] P. Kartsonas. *Praktiki Methodos Ekmathisis tis Psaltikis Tehnis [Practice Methods for Learning Byzantine Chant] [Audio CD]*. The Hunt of St. George Mount Athos.

[10] G. K. Koduri, J. Serrà, and X. Serra. Characterization of intonation in carnatic music by parametrizing pitch histograms. In *Proceedings of ISMIR*, pages 199–204, 2012.

[11] K. Levy and C. Troelsgå rd. *Byzantine chant*. Oxford University Press, accessed August 24, 2011, http://www.oxfordmusiconline.com/subscriber/article/grove/music/04494, 2011.

[12] M. Mavroeidis. *Oi mousikoi tropoi stin Anatoliki Mesogeio [The musical modes in East Mediterranean]*. Fagotto, Athens, 1999.

[13] D. Moelants, O. Cornelis, and M. Leman. Exploring African Tone Scales. In *Proceedings of ISMIR*, pages 489–494, 2009.

[14] H. Purwins, B. Blankertz, and K. Obermayer. A new method for tracking modulations in tonal music in audio data format. In *Proceedings of IJCNN*, volume 6, pages 270–275, 2000.

[15] J. Serrà, G. K. Koduri, M. Miron, and X. Serra. Assessing the tuning of sung indian classical music. In *Proceedings of ISMIR*, pages 157–162, 2011.

[16] A. Thoukididi. *Kripida Byzantinis mousikis [Foundations of Byzantine music]*. Kykkos Church, Cyprus, 2003.

[17] G. Tzanetakis, A. Ermolinskyi, and P. Cook. Pitch Histograms in Audio and Symbolic Music Information Retrieval. *Journal of New Music Research*, 32(2):143–152, 2002.

[18] I. Zannos. Intonation in Theory and Practice of Greek and Turkish Music. *Yearbook For Traditional Music*, 22(1990):42, 1990.
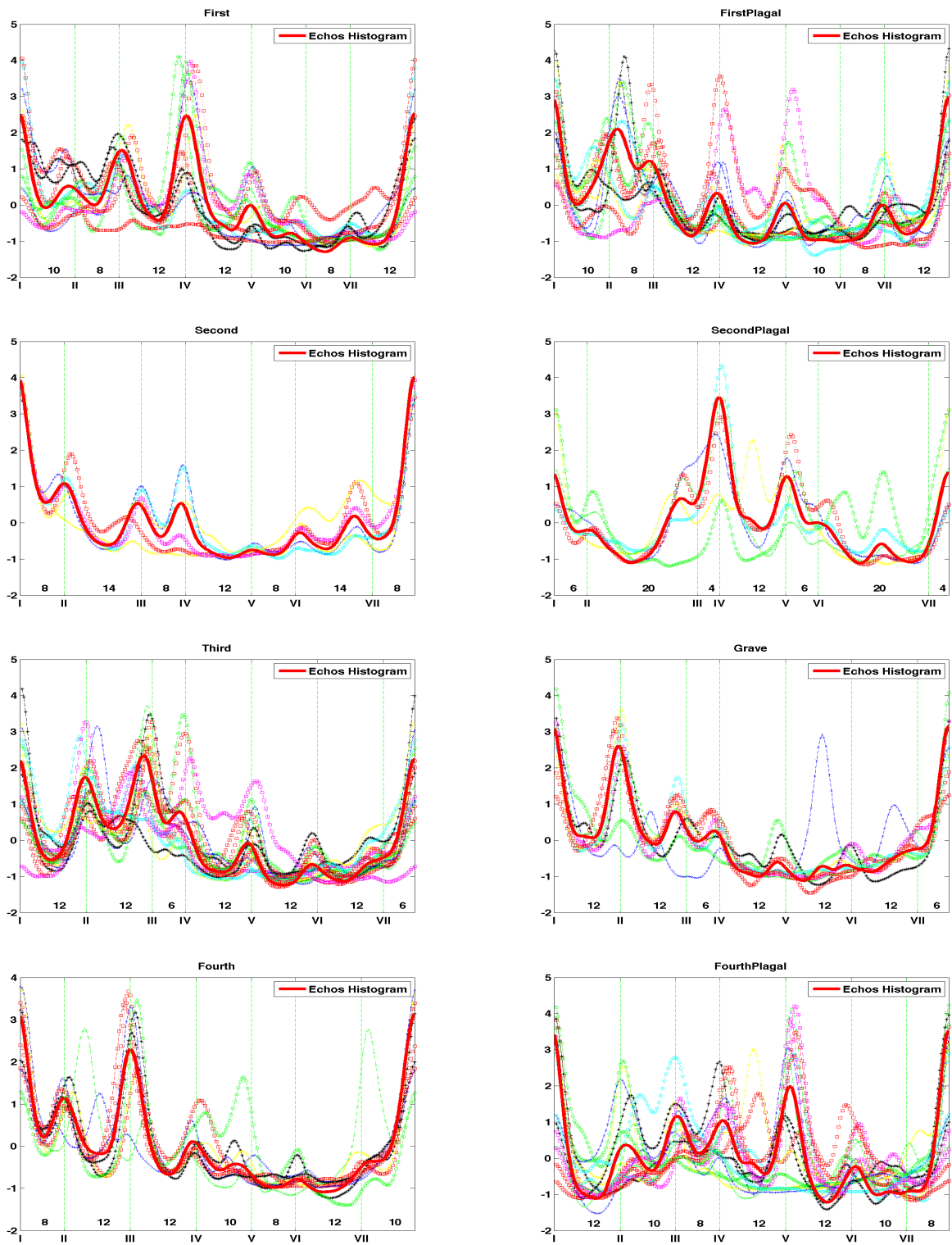
Figure 2. For all recordings of all echoi, the pitch histograms are displayed. The vertical lines indicate pitches of scale degrees according to Chrysanthine theory [16]. The y axis represents the normalized histogram count. The bold red lines represent the echos histogram computed from pitch trajectories across all recordings of the same echos.