On Exploring Vocal Ornamentation in Byzantine Chant

Dimitrios S. Delviniotis and Sergios Theodoridis, *Ilisia, Athens, Greece*

Summary: Objectives. A special vocal ornament in Byzantine chant (BC), the single cycle ornamentation structure (SCOS), is defined and compared with the vibrato with respect to its time (rate, extent) and spectral (slope [SS], relative speaker's formant [SPF] level, formant frequencies [Fi] and bandwidths [Bi], and noise-to-harmonics ratio [NHR]) characteristics.

Study Design. This is a comparative study between the vocal ornaments of SCOS and vibrato, of which time and spectral acoustic parameters were measured, statistically analyzed, and compared.

Methods. From the same hymn recordings chanted by four chanters, the SS, SPF level, FFi, FBi, and NHR difference values between the vocal ornament and its neighbor steady note, and the rate and extent, were compared with those of vibrato.

Results. The mean extent values for SCOS were found to be almost double the corresponding values for vibrato, and the rate of SCOS tends to be different from the rate of vibrato. The difference values of: 1) the NHR, 2) the spectral slope, and 3) the SPF level, between the vocal ornament and its neighbor steady note were found to be: 1) higher for SCOS, 2) mainly lower for SCOS, and 3) lower for SCOS, respectively. No significant differences were detected for the FFi and FBi. The FF1 differences tend to be negative in both ornaments indicating a formant tuning effect.

Conclusions. A new vocal ornament (SCOS) in BC is studied, of which the extent, NHR (HNR), the spectral slope, and the SPF level are different compared to those of vibrato.

Key Words: Singing voice ornaments–Spectral characteristics–Byzantine chant–Formant frequencies–Formant tuning.

INTRODUCTION

In the numerous singing styles found around the world, vocal music involves melodic ornaments that serve to decorate a phrase of melody. Such singing styles are either the "classical" (eg, Western operatic, baroque) or the nonclassical, which include, among others, popular singing styles "contemporary commercial music" (eg, rock, jazz, musicals) and traditional singing style (eg, yodel, overtone singing). In the Western singing style, the vibrato, trill, and trillo are some examples of the ornaments that have been studied.[1–6]

Rate and extent

Vibrato, trill, and trillo consist of a modulation of the phonation frequency (F0) characterized by its oscillation rate and extent. The latter reflects the limits within which the F0 (fundamental frequency) varies. In other singing styles, the vibrato may exist and be produced in a different fashion for each genre of the singing voice, although its presence is a determinant in the characterization of the particular genre.[7] Until today, to our knowledge, the trill and trillo have only few studies for either "classical"[4,5] or nonclassical singing style.[8–10] The vibrato is a natural phenomenon, of which the origin is in a neural center, and it can be caused by tremor in any of the muscles involved; the breathing muscles or the laryngeal musculature.[2,6(p165–170)] The pitch perceived from a vibrato tone of which the rate does not go below a critical value of about 5 Hz is quite stable.[6(p171)] If slower than 5 Hz, the hearing system tends to track an undulating pitch rather than a stable one. The trill is a musical ornament corresponding to a quick alternation between two adjacent scale tones, and the trillo is described as the production of rapid reiterations of a single note.[11] The number of oscillations of vibrato, for a sung note, varies from a few cycles in short sung notes up to many cycles in long ones. In special cases of vocal music ornamentation, such as in coloratura singing, within a fast movement of melody, there is a rapid succession of short notes of a duration comparable with half or one vibrato cycle.[12] The pitch perceived from such a very short note was found to deviate by a decade of cents from the mean frequency value of the note.[13]

The normal rate of vibrato was found to be between 5 and 8 cycles per second (Hz), and the acceptable extent is generally less than 1 semitone (ST).[1(p33–52)] Sundberg reported that vibrato with rate value less than 5.5 Hz sounds unacceptably low, whereas when exceeding 7.5 Hz, vibrato tends to sound nervous.[6(p164)] The vibrato extent can be extended up to 2 ST, whereas for values more than 2 ST, the vibrato tends to sound bad. For early music singers, the average rate values of vibrato, exaggerated vibrato (the vibrato in which the singer allows to extend the frequency range as much as possible), and the trill were found to be inside the range of normal vibrato, whereas the trillo was found to be either slow (2–6.9 Hz) or fast (7.5–12.4 Hz).[4] The average extent values of vibrato, exaggerated vibrato, trill, and trillo for this singing style amount to 1 ST, 2.2 ST, 2.7 ST, and 1.6 ST, respectively.[5,6]

Spectral parameters

For better understanding of the voice functioning, apart from rate and extent, spectral characteristics of the voice signal can be measured to examine both the function of the glottal source and the vocal tract. More specifically, the spectral slope (SS)

Accepted for publication October 18, 2017.
From the Department of Informatics and Telecommunications, University of Athens, Ilisia, Athens, Greece.
Address correspondence and reprint requests to Dimitrios Delviniotis, Department of Informatics and Telecommunications, University of Athens, Ilisia, GR-15784 Athens, Greece. E-mail: ddelvis@di.uoa.gr
Journal of Voice, Vol. 33, No. 2, pp. 256.e17–256.e34
0892-1997

https://doi.org/10.1016/j.jvoice.2017.10.016

and the formants are important components that are related to both the glottal operation and the whole configuration of the vocal tract.[14] The long-time average spectra (LTAS) show characteristics of both the glottis and the vocal tract in speech[15] and singing.[16] The LTAS is a frequency representation of the audio signal providing the average sound level along the frequency axis over a period of time. SS has been shown to be related to the glottal closure during phonation and, more specifically, to the rate of glottal adduction.[17,18] A reduction in SS from a normal value is known to represent hypoadduction, whereas an increased SS is related to hyperadduction.[19] However, the SS depends on SPL of the vocal sound because of the nonuniformly of the LTAS contour in the frequency range when SPL changes.[18] The spectral distribution of power or energy of the vocal sound is related to SS (spectral tilt) and more specifically to the power or energy ratio between the low frequency region, 0–2 kHz, and the high frequency region, 2–4 kHz.[19] As the first formant is located inside this low frequency region, the most of the signal energy is concentrated there. Both the power and the energy ratios have been used for the quantitative evaluation of the voice quality.[20,21] The relative power level of the speaker's formant (SPF) provides a measure of the higher frequency range (2.2–3.7 kHz) contribution to the voice quality relative to a broader frequency range of 2–4.6 kHz. The SPF, which may exist in singing voices,[22] is a single peak between 3 and 4 kHz and it is enhanced in singing due to both the rise of the SS and, possibly, the reasons not sufficiently known yet. Given that the higher formants (F3, F4, and F5) are related to the larynx and pharynx tube configuration,[23,24] the SPF has to do with the characteristics of individual chanter rather than a systematic vocal functioning in BC. In the operatic male singing, because of a clustering of the higher formants, a single peak appears in the spectrum around 2.8 kHz, which is called as singer's formant (SF). In addition to the SS and SPF, the noise-to-harmonics ratio (NHR) is a spectral parameter that measures the hoarseness of a voice and reflects a source of noise in vocal folds or near them, generating aperiodicity of the audio signal.[25] Harmonics-to-noise ratio (HNR) is related to NHR and it is defined as the ratio of the power of the harmonics of the voice signal to the power of the inharmonic content (nonperiodic part of the signal, considered as noise) of the voice signal.[25,26]

Formant tuning

Formant tuning or matching is a central issue in research on both classical and nonclassical singing and it refers to the fact that the singers tune a formant frequency to the frequency of a partial. Soprano singers tune F1 to a frequency close to the first harmonic, in cases where otherwise the fundamental frequency f0 would be higher than F1.[27] The term "formant tuning" is generally used for cases where one of the lowest formant frequencies coincides with the frequency of a source spectrum partial; however, it is a contradictory issue among the researchers. Some authors claim that this effect is a goal of classical opera voice training,[28,29] whereas others claim that none of the two first formants tend to change systematically between scale notes such that either one or both formants to coincide with a spectrum partial.[30]

Summary of previous studies

A summary of methods of analysis in previous works on vocal music ornamentation is presented in Table 1. Vocal music ornamentation has been analyzed, from an acoustical point of view, since the late 1920s when Carl Seashore and his associates exhaustively researched the vocal vibrato and other ornamentation.[1] Vocal vibrato was studied by analyzing the pitch contour resulted from the time-frequency representation of the audio signal either as the fundamental frequency (F0) track[3,5] or as the track of one partial.[31,32] Through locating the troughs and crests at the pitch contour, the main vibrato characteristics of rate and extent signals were estimated in the most of the singing studies. Besides rate and extent, the intonation signal of vibrato was calculated by averaging the extreme values of subsequent vibrato cycles.[31] A more detailed rate, extent, and intonation signals were estimated from a complex signal called "analytic signal," which results from the vibrato signal.[32] Until now, few studies about trills and/or trillo were implemented in which the above parameters of rate and extent were measured from spectrograms.[4,5,9,10]

Some results from previous research while comparing various singing styles with others are considered here. For the Peking opera, it was found that the vibrato rate can be equal to the lowest value of 3.5 Hz,[33] whereas the mean SS in LTAS can be as much as −11 dB/octave in the 0.7–6 kHz frequency band. In the 0.8–1.6 kHz frequency band, the LTAS curve for the Broadway style is about 5 dB higher than the operatic style, whereas the overall slope in the range 0.5–1.6 kHz in speech was rather similar to that demonstrated in the Broadway style.[34] The SS in terms of the singing power ratio between the highest power level of the two frequency bands was for sung vowels significantly greater than that of spoken vowels.[20] By assessing the areas under the LTAS curve for the two frequency bands instead of the highest power level, the power ratio is replaced by the energy ratio. The energy ratio increases while changing the voice "projection" of a singer due to the shift in acoustic energy to the high frequency band.[21] The higher formants are clustered, forming an SF in opera but not in the Broadway style for male singing,[34] whereas for the jazz style, there was a clustering of resonances around 3–3.5 kHz, forming a less prominent peak than that in opera style.[35] Voice perturbation measures were measured for various singing styles,[36] and the HNR was found to be similar for Fado and Western classical singers, and significantly low in comparison to opera, soul, country, jazz, and music theater.[37]

Byzantine chant

Byzantine chant (BC—the chant of the Greek Orthodox Church) is the religious part of what is commonly described by the term "Byzantine Music" which is associated with the medieval sacred chant of Christian Churches following the Constantinopolitan Rite. BC has received little attention compared with the Western styles. Nowadays, as this music has influenced most of the singing types of the Greek music, there is clearly a need for a thorough

TABLE 1.
Summary Reports of Previous Studies on Vocal Ornamentation Analysis

Authors	Singing Style	Type of Ornamentation	Measured Parameter/ Analysis Method or Software	Observations
Jean Hakes, Thomas Shipp, and Thomas E. Doherty	Early music	Vibrato—exaggerated vibrato—trill—trillo	Rate—extent—oscillation jitter (rate—extent)/Zero crossings for F0 calculation—manually	Vibrato, exaggerated vibrato, and trill rates were similar, but trillo rate was much faster. Average extent of oscillation for vibrato is 1 ST, for exaggerated vibrato 2.21 ST, for trill 2.71 ST, and for trillo 1.64 ST.
Jean Hakes, Thomas E. Doherty, and Thomas Shipp	Early music	Trillo	Rate/Zero crossings for F0 calculation—manually	Continuous trill at all pitch levels formed two distinct clusters: slow (2–6.9 Hz) and fast rates (7.5–12.4 Hz)
Eric Prame	Western classical	Vibrato	Extent/time spectrograms—F0 at all crests and troughs on a distinct partial	Mean vibrato extent for individual tones ranged from ±34 cents to ±123 cents/mean across tones and singers amounted ±71 cents. Extent is negatively correlated with tone duration and positively correlated with intonation.
Ixone Arroabarren, Miroslav Zivanovic, Jose Bretos, Amaya Ezcurra, and Alfonso Carlosena	Western classical	Vibrato	Rate, extent, and intonation/band pass filtering and analytic signal's estimation (decomposition into instantaneous amplitude and frequency)	Three main time-dependent functions are proposed and associated to acoustically defined parameters: rate, extent, and intonation. Significant differences in the vibrato among different singers. Good agreement with perception.
Ana P. Mendes, Aira F. Rodrigues, and David Michael Guerreiro	Fado singing (Portugal)	Sustaining vowels—vibrato	Jitter, shimmer, NHR, LTAS/ KayPENTAX software	Jitter and shimmer mean values were higher for singers compared with nonsingers, whereas NHR was similar. Jitter was higher compared with country, musical theater (MT), soul, and jazz singers and lower than pop and Western classical. NHR was similar for Western classical. No singer's formant. Male rate (5.72 ± 0.72), extent (2.39 ± 0.75).
Ananya Bonjyotsna and Manabendra Bhuyan	Indian pop singing	Vibrato—mordent	Rate and extent of vibrato—Time duration and extent of mordent	The rate and extent values ranged 5.26–6.06 Hz and 41.64–66.79 cents, respectively. Vibrato rate depends on the tempo of the melody. Time duration of the mordent range: 0.1–0.2 s and its extent ranged from 1 to 38 Hz (about 1.5 ST).
Johan Sundberg, Lide Gu, Qiang Huang, and Ping Huang	Peking opera	Vibrato—nonvibrato	Vibrato rate—spectral slope—singer's formant/LTAS	Vibrato rate was about 3.5 Hz. LTAS did not show a singer's formant with mean slope in range 700–6000 Hz of −11 dB/octave in singing
Johan Sundberg, Fillipa M. B. La, and Brian P. Gill	Classical and nonclassical—musical theater	Vibrato	Formant frequencies and bandwidths/ Inverse filtering—DeCap	A rising spectrum envelope over the three lowest partials on the highest tones in classical style. F1 coincided with H2 for classical and nonclassical styles. No nonlinear source-filter interaction in the formant tuning situation, for classical style, whereas for nonclassical one, irregularities were found.
Caitlin J. Butte, Yu Zhang, Huangqiang Song, and Jack J. Jiang	Jazz—opera—soul—country—musical theater—pop	Sustained vowels	Jitter, shimmer, signal-to-noise ratio/Cspeech	Jitter and shimmer ranges of 0.3–1.6% and 0.7–6.0% for normal speaking, respectively. The pop style had mean jitter of 1.13% and shimmer of 6.78%. Jazz, soul, MT, and country had normal values in jitter and shimmer. The opera style had normal jitter (0.52%) but high median shimmer values (7.07%). Normal values for SNR are in the range of 9–30 dB. All styles into consideration fell into normal range. Pop style had low median SNR (12.3 dB) in comparison with the others (17.1–24.4 dB).
Claudia Manfredi, Davedi Barbagallo, Giovanna Baracca, Silvia Orlandi, Andrea Bandini, and Philippe H. Dejionckere	Classical—Jazz	Sustained notes—vibrato	Vibrato rate, extent, duration, regularity—Singer's formant/ BioVoice	Vibrato rates did not differ among the two styles. Extent was significantly larger in classic singers, particularly classic females. The novel parameters of vibrato jitter and shimmer accounting for the vibrato regularity were significantly better in classic singers. The duration of vibrato was significantly larger in classic singers. The clustering of resonances around 3–3.5 kHz was significantly stronger in classic singers and the SF was typical for males.

study. Some of the vocal ornaments met in BC have, still, survived in the traditional Greek music such as folk songs, popular songs, and rebetika songs.[38] Recently, from the previously mentioned vocal ornaments, the BC vibrato was analyzed,[39] with typical rate values of 5.13 ± 0.96 Hz, which are slightly lower than the typical values in the Western opera, and the average extent values were as much as 0.5 ± 0.2 ST. Yet, in BC, the formant tuning effect was recently studied, and it was reported that F1 and/or F2 are close to a partial, irrespective of whether this coincidence is done in a systematic way or not.[40] Finally, apart from vibrato, special vocal ornaments of one vibrato cycle in time length are commonly used by the BC performers (chanters). The waveform of their F0 modulation in most cases resembles a single period pseudo-sinusoidal with initial phase $-\pi/2$ and a positive offset level relative to the stable part of its neighbor note. This ornament has not been specified yet in a systematic way in BC and a thorough study is needed for that. As this ornament appears in isolation in time, we call it as the "single cycle ornamentation scheme (or structure) (SCOS)" and its neighbor stable note is located in time prior or after that or both (Figure 1).

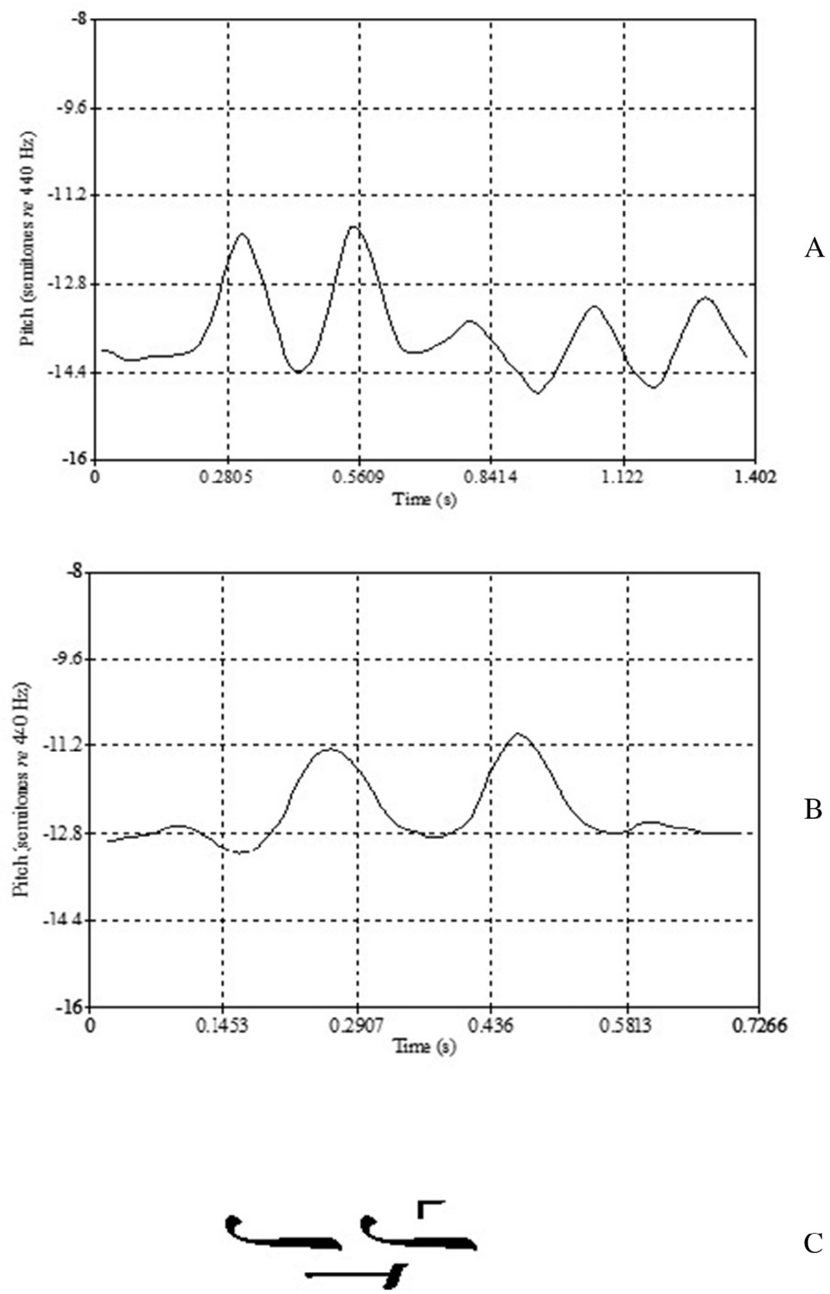

FIGURE 1. **A.** Pitch track of two single cycle ornamentation structures (SCOS) in succession followed by a vibrato on the same note. **B.** Pitch track representation of the two SCOS separated to each other by a steady part between them. **C.** Byzantine chant (BC) musical score of BC phrase frequently appeared in BC melodies, which corresponds to the pitch track in (**B**).

The main goal in this study was to compare the SCOS with the vibrato vocal ornament with respect to the time characteristics of the rate and extent, and the spectral characteristics of the NHR, LTAS slope, and SPF level; the vibrato was considered here as a reference ornamentation structure. To this end, (1) we introduce the concept of the SCOS, (2) we specify the range of the acoustical parameter values for both the vibrato and the SCOS, and (3) by measuring and statistically processing the above acoustical parameters, for the SCOS and vibrato, we test the assumption that the SCOS is different from vibrato, that is, the SCOS is not simply a single cycle vibrato.

METHODS

The single cycle (F0 modulation) ornamentation structure (SCOS)

As the SCOS has not been defined in the past, an exploratory research design was first applied in this study. By examining many pitch tracks from BC melodies, performed by a number of famous chanters, it was observed that there were many pitch track structures, which can be considered as sinusoidal modulations of the fundamental frequency. Apart from the vibrato on one note that is already known to the researchers, other pitch track structures do also exist. Such structures either appear between successive notes or decorate the notes themselves (Figure 1A—sound clip). In the cases where such a structure exists between successive notes, as in a virtuoso passage, this is hard to be studied because of their unknown pitches.[13] On the contrary, when such a structure "rides" a note with specified pitch, that structure can be analyzed by comparing it with the steady part of the note over which it is located. By examining many BC musical phrases that included pitch structures, as the previously reported, with specified pitches while taking into account their corresponding music score, we realized that such musical phrases often correspond to the musical score structure as that shown in Figure 1C.

For this BC score musical phrase, BC theoreticians report that it is chanted with a special vocal "roughness" or with a "vocal undulation," meaning a special way of the artistic performing with abrupt pitch changes.[41(p58)] When looking at the pitch track of this musical phrase, we see that one or two simple pitch structures exist, with each one consisting of a fast rise and fall of the F0 of the note which these structures decorate (Figure 1). This note has a certain pitch and this ornamentation structure is frequently found in BC melodies. It is essentially a rapid change between the modulated note and a higher note with perceptually unspecified pitch. For the previous musical phrase, this is performed with two successive SCOS on its pitch track, separated to each other by a part of stable pitch (Figure 1B). The time instants at the local minima on both sides of SCOS are considered as its time borders (Figure 2). However, in rare cases, these minima disappear and the SCOS tends to end on both sides with exponentials or hyperbolics. In such cases, the SCOS time borders are specified as the time instants at which

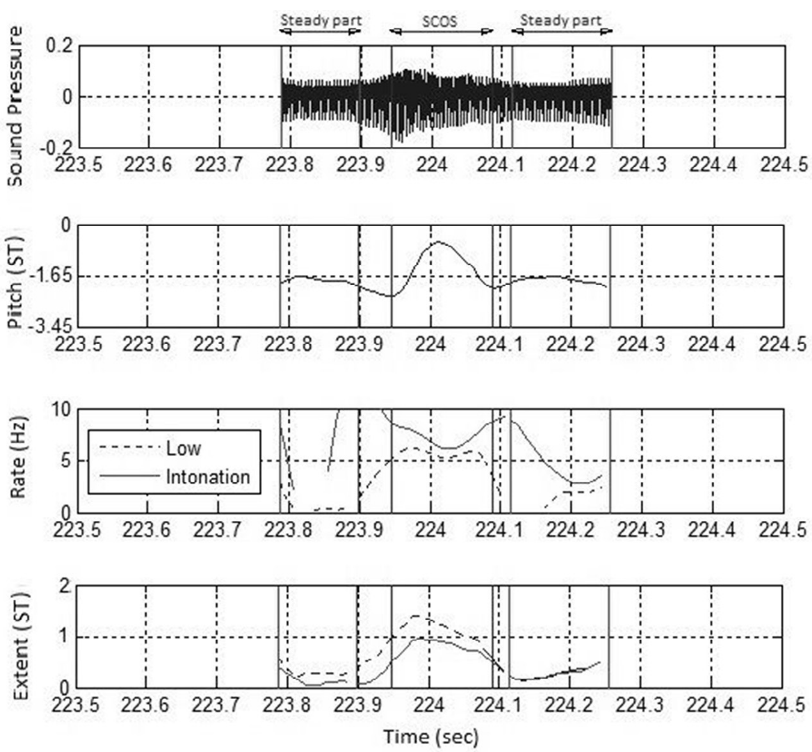

FIGURE 2. Typical example of a single cycle ornamentation structure (SCOS) with its time borders and the steady parts of the note carrying it. The time instants of local minima that are close to the peak of the SCOS are considered as the SCOS time borders. Each steady part must have standard deviation of its pitch values smaller than 1/9 semitones (ST). Corresponding rate and extent trajectories were calculated based on intonation track's estimation through low filtering of the audio signal (dotted line) and through cubic interpolation of the half-way distance points between the extremes of the pitch track (solid line).

the pitch variation becomes smaller than 1/9 ST; a value that is the upper limit for each steady note close to the SCOS to be considered as a "stable" pitch note. In these rare cases, the SCOS deviates from a sinusoidal, the SCOS time length is widened, and the mean values of the rate and extent appear to be somewhat lower.

Because of the improvisation of which BC performance is generally characterized, it is possible that if performing a particular musical score, phrase twice this does not correspond to the same pitch track representation even if performed by the same BC chanter. Provided that the SCOS samples can be searched within the music score phrases resembling to that in Figure 1C, they could reliably be identified from the music score and then isolated from the respective pitch track.

Method's design

From experience, it is known that the note, over which the SCOS exists, has a clearly heard pitch that is not affected by the SCOS existence. In other words, the steady part of that note somewhat dominates the pitch perception of the whole music structure of SCOS—steady note. In light of this, the SCOS was studied in reference to the steady part of the carrying note, and the SCOS acoustical parameters were compared with the respective parameters of the steady part of the carrying note. The idea for that comparison was that the configuration of the vocal system when performing an SCOS on a note and its time neighboring steady part remains unchanged if no difference between SCOS and steady part exists. Hence, any variations in the acoustical parameter values also reflect variations in the voice functioning between SCOS and steady part. For this purpose, each sample of SCOS for analysis must be accompanied by a steady part for comparison; yet, both SCOS and steady part must correspond to the same vowel and intensity.

Data selection

The SCOS excerpts were extracted from the sound signal based on the following criteria: (1) a SCOS coexists along with a note of constant pitch (steady note with standard deviation smaller than 1/9 ST) and it is clearly heard, (2) the SCOS time borders are located, as mentioned previously (at the local minima in both sides of the SCOS), and (3) the vowel of both the SCOS and its steady part is the same. Also, the vibrato notes were extracted in a similar way; namely, they were accompanied by a note with constant pitch and with the same pitch and vowel as the vibrato note.

To identify such pitch track structures, the most suitable melody in BC is the Papadic[42]-type melody in BC, which is featured by long musical phrases with very slow notes on a single vowel (0.75 second per note), from one side, and by embellishment structures containing SCOSs accompanied by neighbor steady notes from the other side. Also, vibrato notes, along with steady notes of the same pitch and vowel, can also be found within this type of melody.

Participants

Four male BC professionals participated as subjects. The selection criteria for their participation were (1) to be experts in BC performance, (2) to have healthy and musically "flexible" voices, and (3) each one performs in a different style of chanting than the others. All of them chanted, in whatever tone was convenient for them (though the tonic notes differ to each other no more than three ST), the same cherubic hymn[43] belonging to Papadic type of melody, in a normal loudness. The tempo was in the range of 75–88 beats/min. To ensure that each chanter uses comfortable phonation, he or she should sing with effortless voice, within an octave. By instructing the chanters to chant around the centers of their voice ranges, their voices operated within the boundaries of a convenient way of phonation. Furthermore, the F0 range of the BC melody should be contained in the F0 vocal range for each one of the participants. As in a BC melody, the most frequent note is the tonic note in which the BC mode is based on, a suitable BC mode was selected such that the tonic note lies in the middle of the chanter's voice range. Hence, it is reasonable to infer that most of the SCOS pitch track structures were produced in the middle of the voice range of a particular BC performer because most of the notes lied around the center of his voice range. By choosing the BC fourth main mode in a Papadic melody, which has a tonic note located at the middle of F0 (around G3) range, the above requirements were fulfilled and a large number of SCOS were made available for the analysis.

Recording procedure

The audio signal was recorded through a RODE K2 condenser microphone (RØDE, Sydney, Australia) located at 30 cm from the subject. This signal was digitized at a rate of 44,100 Hz, 16 bit depth, pulse-code modulated, and recorded as a wav file on Yamaha AW 16G Digital Audio Workstation (Manufactured by Yamaha Corporation, Taiwan). These recordings are part of the tagged Damaskinos prototype acoustic corpus of Byzantine ecclesiastic chant voices.[44] The microphone was calibrated for absolute SPL measurements by recording a sample of singing at the same time and distance with the wav file via a Brüel and Kjær Integrating Precision Sound Level Meter, type 2230 (Brüel & Kjær, Denmark). The average SPL of this sample was announced in the recording file with the microphone distance. Recordings were implemented in a "quiet" room, volume of about 4 m^3, with reverberation time less than 0.1 second. The signal-to-noise ratio for all the recordings was equal to 41.3 dB on average, and the mean absorption coefficient for the room was about 0.92.

Acoustical analysis

Audio files were analyzed using the software package for voice analysis, *PRAAT*,[45] and the data were further analyzed in the *MATLAB*[46] programming platform for the final results. Both F0 and intensity were measured with time step and window of 10 milliseconds, and both the formant frequencies and the bandwidths were calculated through an iterative algorithm refining the final values through LPC (robust method), with time step of 10 milliseconds, window of 25 milliseconds and a pre-emphasis filter for above 50 Hz.[47] The number of standard deviations, iterations and the tolerance were selected to be 1.5, 5 and 0.000001, respectively. LTAS were calculated from the

TABLE 2.
Number of SCOS and Vibrato Samples for Each Chanter, Vowel, and Pitch Range of Each Chanter

		Chanter 1		Chanter 2		Chanter 3		Chanter 4	
Micromelody	SCOS		171		306		104		94
	Vibrato		158		18		137		105
Vowel5	a		79		74		56		44
	e		89		90		65		56
	i		87		68		76		53
	o		61		78		33		35
	u		13		14		11		11
Pitch ranges	1st	Do3—Re3#	51	Si2—Re3	16	Re3—Fa3	20	Do3—Re3#	19
	2nd	Re3#—Fa3#	78	Re3—Fa3	89	Fa3—Sol3#	61	Re3#—Fa3#	53
	3rd	Fa3#—La3#	173	Fa3—La3	186	Sol3#—Do4	132	Fa3#—La3#	107
	4th	La3#—Do4#	27	La3—Do4	33	Do4—Re4#	28	La3#—Do4#	20

.*Notes:* The number of samples for each vowel and chanter is large enough for analysis, except for the vibrato of chanter 2 and the vowel /u/ for all chanters. Also, the chanters have almost identical pitch ranges.

normalized audio signals in the frequency band [0–5000 Hz], with a resolution bandwidth of 100 Hz. The term LTAS is not precise here because of the small time length of the analyzed audio samples; however, this "LTAS" is in fact a spectral representation of the audio signal describing its power distribution in the frequency zones (bins). LTAS slope was calculated as the energy ratio between the frequency ranges of 0–1 and 1–4 kHz[48] and the relative level of SPF between the regions of 2.2–3.7 and 2–4.6 kHz.[49] The boundary of 1 kHz was determined because in BC, the largest percentage of voice power was found within the 0 and 1 kHz frequencies.[50] Instead of considering the highest peak in the 0–1 kHz and 1–4 kHz bands, the above energy ratio was calculated in terms of the average power spectral density for each frequency band. However, the average power spectral density was calculated after it was converted from logarithmic to linear values ("Energy" method).[45]

Glottal noise measures NHR[51] and HNR[a,7,52(p99)] were calculated through the voice report in the PRAAT software,[45] whereas the rate and extent of both SCOS and vibrato were calculated through a MATLAB software based on the Hilbert transform of the pitch track signal after subtracting its intonation track.[39] According to MATLAB software, the intonation track of the pitch signal is estimated either through low-pass filtering of the initial pitch signal (first method) or through cubic interpolation of the half-way distance points between the extremes of the pitch track (second method).

To minimize the dependency on vowel and pitch, for all the spectral parameters, the difference in values between the vocal ornament (SCOS or vibrato) and the steady note connected to that vocal ornament was calculated. It has been hypothesized that the sound level does not change between the vocal ornament and the steady note. For each parameter, the difference values and the percentage differences were calculated according to Equations (1) and (2), respectively,

$$\Delta parameter = parameter_{micromelody} - parameter_{steady\ note}, \quad (1)$$

$$\Delta parameter_{percentage} = 100 * \frac{parameter_{micromelody} - parameter_{steady\ note}}{|parameter_{steady\ note}|}, \quad (2)$$

where the *parameter* stands for F_i, B_i, HNR, NHR, SlopeE, and SFheightE, which are presented next. For the F_i and B_i, only the nonpercentage differences were calculated (Equation 1). The suffix "E" for the SS and SPF relative level (SFheight) stands for the measured values through "energy" method (E). All the difference values are represented with variables with the same parameter name but with Greek letter Δ as prefix.

Statistical analysis

Statistical correlations through independent samples *t* tests between SCOS and vibrato were applied on the acoustic parameters of the rate and extent. For the rest of the parameters F_i, B_i, HNR, NHR, SS, and SPF's relative height, both one sample *t* tests and independent samples *t* tests between SCOS and vibrato were applied on the calculated parameters of ΔHNR, ΔNHR, ΔSlopeE, ΔSFheightE, ΔF_i, and ΔB_i ($i = 1$–5), as well as on the percentage counterparts of ΔHNR%, ΔNHR%, ΔSlopeE%, and ΔSFheightE% (Equation 2). To confirm the results from the previous statistical analysis, a univariate analysis of the variance tested the effects of vocal ornament type, vowel, and pitch on the rate and extent, and on the NHR, HNR, SS, and SPF's height differences. To examine whether a formant tuning effect exists, one sample *t* test was applied on the percentage difference between the first formant frequency and the second or the third harmonic frequency of the voice spectrum.

RESULTS

The number of the sample excerpts for both SCOS and vibrato per vowel, chanter, and its pitch range were shown in Table 2. It must be noted that for chanter 2, the number of

[a]The harmonics-to-noise ratio (HNR) is a measure that quantifies the dominance of the periodic part level over the aperiodic one, expressed in dB units. The noise-to-harmonics ratio (NHR) is the average ratio between noise in the frequency band of 1.5–4.5 kHz and the harmonic energy in the frequency band of 0.07–4.5 kHz: it is sometimes referred to as a low-frequency HNR.

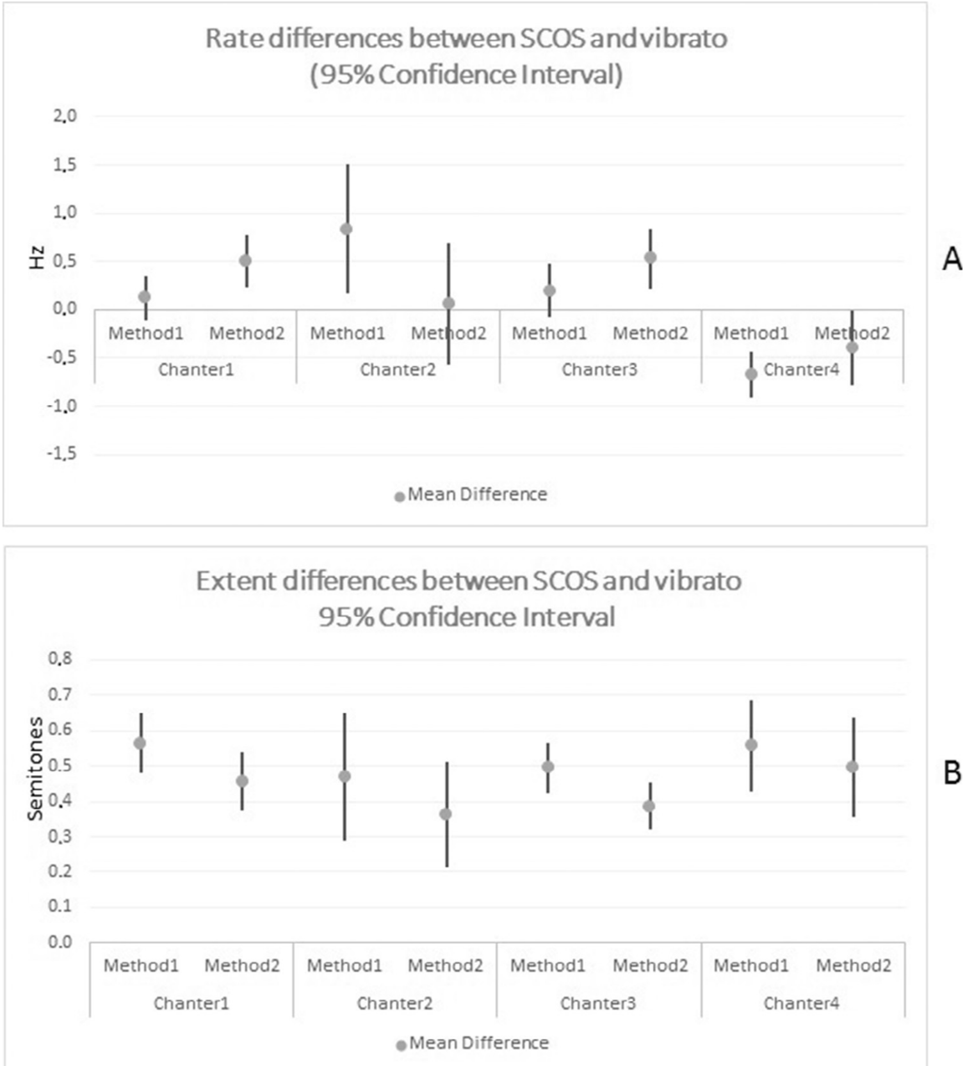

FIGURE 3. Mean rate differences between single cycle ornamentation structure (SCOS) and vibrato for each chanter and both the methods of measurements (**A**). From these differences, the statistically significant ones, although statistically significant differences, although not for both methods, are positive for chanters 1, 2, and 3, and negative for chanter 4. Mean extent differences between SCOS and vibrato for each chanter and both methods of measurements (**B**). All the extent differences are positive and statistically significant.

vibrato samples was much smaller compared with that of SCOS, indicating that the vibrato effect was a rather rare ornament for that chanter. In spite of this, the number of samples for each vowel and chanter was large enough for statistical analysis, except for the vibrato samples of chanter 2 and the vowel /u/ for all the chanters. The pitches of the samples varied inside the same frequency range as the chanters have almost identical pitch ranges (Table 2).

Rate and extent of SCOS and vibrato

An equal variances *t* test with significance level of $\alpha = 0.05$ (and for all the subsequent *t* tests) revealed a statistically reliable difference between the mean values of the extent for SCOS and vibrato, for all four chanters and the two methods of measurements; the SCOS extent was found to be higher than that of vibrato for all chanters and methods of measurement (Figure 3B, Table 2). The mean extent values for SCOS tend to be twice the respective values for vibrato (Table 3A). A similar *t* test for the rate also revealed statistically reliable differences between the mean rate values of SCOS and vibrato for all the chanters, from one side, but for different method, from the other side (Table 3B). Also, concerning the rate, the statistically significant difference between SCOS and vibrato did not have the same sign for all the above four chanters; for chanter 1, chanter 2, and chanter 3, this difference was positive (6.32–5.82 Hz, 5.86–5.02 Hz, 6.35–5.82 Hz), whereas for chanter 4, it was negative (4.01–4.68 Hz, 4.67–5.07 Hz) (Table 3A, Figure 3A). The average of the means across all four chanters was different for each one of the methods, namely, the first method yielded larger values than the second one for the extent of both SCOS and vibrato, whereas for the rate, the second method produced larger values than the first one (Table 3A).

TABLE 3.
(A) Mean and Standard Deviation Values of the Rate (Hz) and Extent (ST) of the SCOS and Vibrato for All Four Chanters and the Two Methods of Measurements* and (B) Equal Variances *t* Test Results for the Rate and Extent Between the Mean Values of SCOS and Vibrato for Each Chanter and Method of the Analysis†

A

		Chanter 1		Chanter 2		Chanter 3		Chanter 4			
Parameter/Method	Micromelody	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Average mean	Average SD
Rate1‡ (Hz)	SCOS	5.28	1.12	5.86	1.34	5.27	1.10	4.01	0.75	5.11	1.08
	Vibrato	5.16	0.86	5.02	1.73	5.08	0.95	4.68	0.77	4.98	1.08
Rate2 (Hz)	SCOS	6.32	1.42	7.16	1.28	6.35	1.23	4.67	1.44	6.13	1.34
	Vibrato	5.82	0.82	7.10	1.19	5.82	1.09	5.07	1.10	5.95	1.05
Extent1 (ST)	SCOS	1.21	0.48	1.10	0.37	1.09	0.32	1.22	0.52	1.15	0.42
	Vibrato	0.64	0.20	0.63	0.30	0.60	0.16	0.66	0.32	0.63	0.24
Extent2 (ST)	SCOS	1.03	0.46	0.87	0.30	0.91	0.29	1.09	0.59	0.98	0.41
	Vibrato	0.58	0.20	0.50	0.33	0.52	0.16	0.60	0.31	0.55	0.25

B

	t Test for Equality of Means							
	Chanter 1		Chanter 2		Chanter 3		Chanter 4	
Parameter/Method	df	Sig. (2-tailed)	df	Sig. (2-tailed)	df	Sig. (2-tailed)	df	Sig. (2-tailed)
Rate1‡	317.216	0.268	322	**0.012**	239	0.138	193	**0.000**
Rate2	263,031	**0.000**	312	0.838	239	**0.000**	190	**0.032**
Extent1	229.408	**0.000**	322	**0.000**	140.471	**0.000**	151.277	**0.000**
Extent2	225.919	**0.000**	312	**0.000**	147.012	**0.000**	134.880	**0.000**

Notes: Significance level of $\alpha = 0.005$.
* The extent values for SCOS are almost twice that for vibrato, for both methods.
† The bold characters indicate the statistically important differences between SCOS and vibrato. The extent differences are statistically important for all three chanters and both methods.
‡ Rate1 (first method), Rate2 (second method), Extent1 (first method), and Extent2 (second method).
Abbreviations: df, freedom degrees; SCOS, single cycle ornamentation structure; SD, standard deviation; Sig., sigma value; ST, semitone,

HNR and NHR for SCOS and vibrato

The HNR and NHR parameters, representing the ratio of energies of the harmonic and nonharmonic (noise) contents of the signal, are measured in decibels and in pure decimals, respectively. Through a one-sample *t* test ($\alpha = 0.05$) for the ΔHNR, ΔNHR, ΔHNR%, and ΔNHR%, all these differences between SCOS or vibrato and steady note were found to be statistically significant (Table 4A—bold characters). An equal variances *t* test ($\alpha = 0.05$) revealed statistically reliable differences between SCOS and vibrato types of vocal ornaments for the mean values of both ΔHNR and ΔNHR values across all four chanters; the ΔNHR value of SCOS was higher than that of vibrato and inversely for the ΔHNR value, for all four chanters (Table 4). Also, the mean percentage of ΔNHR values are higher for SCOS than for vibrato by about two to four times, whereas the mean percentage of ΔHNR are lower for SCOS than for vibrato by about two times on average (Table 4A). These findings mean that the nonharmonic content of the vocal sound when the SCOS ornament is chanted is bigger than this one when the vibrato ornament is chanted; each one in comparison with the inharmonic content of the associated steady note. Also, both SCOS and vibrato are characterized by higher NHR and lower HNR from the steady note.

Spectral slope of SCOS and vibrato

The equal variances *t* test ($\alpha = 0.05$) revealed statistically reliable differences of the mean ΔSlopeE between SCOS and vibrato for chanters 1, 2, and 3 but not for chanter 4 (Table 5B). The same result holds for the percentage values ΔSlopeE% too. A one-sample *t* test for the ΔSlopeE and ΔSlopeE% parameters showed statistically significant differences between the SCOS or vibrato and steady note for the most of the mean values (Table 5A—bold characters). All the mean ΔSlopeE values for SCOS and vibrato were around zero and not higher than 1.5 dB. Also, the mean ΔSlopeE values for the SCOS were lower than those for vibrato for chanter 1, chanter 3, and chanter 4 by 1.33, 1.23, and 0.69 dB, on average, respectively, but for chanter 2, the corresponding average value was higher for SCOS than for vibrato by 1.58 dB (Table 5A, Figure 4). The mean standard deviations between SCOS and vibrato were calculated as 2.60, 2.02, 4.13, and 2.82 dB (Table 5A). The means of the percentage ΔSlopeE% values between SCOS and vibrato differed across the chanters by –25.67, 22.01, –81.34, and –9.34, namely, with similar signs of difference as for ΔSlopeE values (Table 5A). Also, the corresponding mean standard deviations between SCOS and vibrato were calculated as 58.41, 25.37, 198.82, and 39.85 dB (Table 5A).

SPF level in SCOS and vibrato

Based on a one-sample *t* test, also some statistically significant ΔSFheightE values for SCOS or vibrato were found and especially for vibrato (Table 5A—bold characters). Another equal variances *t* test between SCOS and vibrato showed statistically significant differences in the mean values of the ΔSFheightE for only chanter 1 and chanter 3. For the other two chanters, this *t* test failed to reveal statistically reliable differences between SCOS and vibrato (Table 5B). These statistically significant differences showed that the mean values were slightly lower in SCOS than in vibrato, roughly, at most 1 dB: [(−0.31, 0.45 dB—chanter 1), (−0.04, 0.99 dB)—chanter 3] (Table 5A). The corresponding mean values for chanters 2 and 4 for the SCOS were also slightly lower than those for the vibrato [(–2.40, –2.10 dB), (0.15, 0.25 dB)], but these changes could be considered as negligible because of the large standard deviations compared with their means (Table 3A, Figure 5). Statistically significant differences between SCOS and vibrato were found for the percentage mean values of the ΔSFheightE also for chanters 1 and 3 (Table 5B). For chanters 2 and 4, there was a small change in the mean value of ΔSFheightE% between SCOS and vibrato for chanter 2, whereas for chanter 4, the large variability of the values (almost 10 times of the mean value) "blurred" the change of means (Table 5A).

Formant changes in SCOS and vibrato

Formant frequencies and bandwidths were estimated, for all the samples, for both vocal ornaments and the respective neighbor steady notes. Their mean values along with their bandwidths, for all four chanters, each vowel and vocal ornament, are presented in the Supplementary Appendix A (Tables A1, A2). By considering, for simplicity, the division of vowels into back and front ones, we organized the vowels to be analyzed into three groups; the "back" vowels /a,o,u/, the "middle" /e/, and the "front" /i/ vowels, which were considered for further analysis. In such a way, the number of back vowels was large enough for statistical analysis and thus overcoming the problem with the small number of the vowel /u/ samples. An equal variances *t* test was applied to the data where the formant frequencies (F_i) and bandwidths (B_i), and the formant frequencies (ΔF_i) and bandwidths (ΔB_i) differences between vocal ornament and steady note were considered as the independent variables for each of the three previous groups of vowels, whereas the type of vocal ornament was considered as the constant factor (Supplementary Tables A1, A2, B1, B2). For both the formant frequencies and the bandwidths (F_i, B_i), no systematic differences were detected across all four chanters; however, for each chanter, some personalized statistically reliable differences were found: [(vowels /a,o,u/, F3, F4, B3 | vowel /e/, F3, B1, B2, B4, B5 | vowel /i/, F2, F3, F4, B1), for the chanter 1, (vowels /a,o,u/, B5 | vowel /i/, B3, B4), for chanter 2, (vowels /a,o,u/, B4 | vowel /e/, B2 | vowel /i/, F3, F4, B3, B5), for chanter3, (Vowels /a,o,u/, F2, B1, B4 | vowel /e/, F1, F3, F5, B2, B3 | vowel /i/, F1, F2, F3, F4, B1, B2, B4, B5), for chanter 4] (Supplementary Tables A1, A2). Also, for both the formant frequencies and the bandwidth differences (ΔF_i, ΔB_i), no systematic differences of the mean difference values were detected across all four chanters: [(vowels /a,o,u/, ΔF1, ΔB3 | vowel /e/, ΔB2 | vowel /i/, ΔB1), for chanter 1, (vowel /e/, ΔF1 | vowel /i/, ΔF1, ΔB2, ΔB3, ΔB4), for chanter 2, (vowel /i/, ΔF1), for chanter 3, (vowel /i/, ΔF1, ΔF2, ΔF4, ΔB2)], for chanter 4 (Supplementary Tables B1, B2). However, some differences, statistically reliable or not, showed an interesting finding: almost all the ΔF1 mean differences (Supplementary Table B1), although not statistically significant, were negative.

TABLE 4.
(A) Mean and Standard Deviation Values of the ΔNHR and ΔHNR, and Their Percentage Values for the SCOS and Vibrato, and All Four Chanters* and (B) Equal Variances *t* Test Results for the Harmonics-to-Noise Ratio and Noise-to-Harmonics Ratio Between the Mean Values of SCOS and Vibrato for Each Chanter†

A

		Chanter 1		Chanter 2		Chanter 3		Chanter 4	
	Micromelody	Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation
ΔNHR	SCOS	**0.047**	0.038	**0.042**	0.033	**0.045**	0.050	**0.034**	0.032
	Vibrato	**0.009**	0.010	**0.017**	0.030	**0.013**	0.014	**0.011**	0.011
ΔHNR	SCOS	**–9.852**	3.577	**–9.506**	2.928	**–8.546**	4.897	**–8.860**	3.637
	Vibrato	**–4.521**	2.589	**–4.788**	4.011	**–5.364**	3.663	**–5.538**	3.218
ΔNHR%	SCOS	**1441.98**	1344.54	**1230.57**	940.32	**1166.44**	2027.82	**1142.70**	1505.29
	Vibrato	**295.59**	256.28	**624.75**	1114.39	**456.12**	506.97	**458.89**	558.36
ΔHNR%	SCOS	**–37.35**	12.64	**–36.41**	10.62	**–33.13**	19.19	**–33.16**	12.66
	Vibrato	**–16.99**	9.56	**–17.70**	14.87	**–19.78**	13.54	**–20.05**	11.06

B

	t Test for Equality of Means							
	Chanter1		Chanter2		Chanter3		Chanter4	
	df	Sig. (2-tailed)	df	Sig. (2-tailed)	df	Sig. (2-tailed)	df	Sig. (2-tailed)
ΔNHR	193.776	**0.000**	322	**0.000**	183.956	**0.000**	197	**0.000**
ΔHNR	309.642	**0.000**	322	**0.002**	115.120	**0.000**	114.584	**0.000**
ΔNHR%	183.325	**0.000**	322	**0.009**	112.814	**0.001**	115.751	**0.000**
ΔHNR%	314.903	**0.000**	18.035	**0.000**	176.486	**0.000**	197	**0.000**

Notes: Δ prefix: Difference between vocal ornament and steady note value. Significance level of $\alpha = 0.005$.

* For all chanters, the mean ΔNHR values are higher for SCOS than for vibrato and the mean ΔHNR values are lower for SCOS than for vibrato. The bold characters indicate the statistically significant differences between vocal ornament and steady note.

† The bold characters indicate the statistically significant differences between SCOS and vibrato. All the differences between SCOS and vibrato, for all chanters, are statistically significant.

Abbreviations: ΔNHR, noise (nonharmonic) to harmonic ratio; ΔHNR, harmonic to noise (nonharmonic) ratio in decibel units; df, freedom degrees; SCOS, single cycle ornamentation structure; Sig., sigma value.

TABLE 5.
(A) Mean and Standard Deviation Values of the Differences Between Vocal Ornament (SCOS and Vibrato) and Steady Note, for the Spectral Slope and the Spectral Height Relative Level of Speaker's Formant (SPF), for Each Chanter and Vocal Ornament (SCOS and Vibrato)[*] and (B) Equal Variances *t* Test Results for the Slope and SPF Height Level Between the Mean Values of SCOS and Vibrato for each Chanter and Method of the Analysis[†]

A

		Chanter 1		Chanter 2		Chanter 3		Chanter 4	
Parameter	Micromelody	Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation	Mean	Std. Deviation
ΔSlopeE[‡]	SCOS	–1.37	2.42	0.35	1.76	–0.34	4.08	–0.92	2.72
	Vibrato	–0.05	2.78	–1.23	2.28	0.90	4.18	–0.23	2.91
ΔSlopeE%	SCOS	–34.76	70.65	1.40	19.43	–95.47	273.11	–20.55	38.98
	Vibrato	–9.09	46.18	–20.61	31.31	–14.13	124.53	–11.21	40.71
ΔSFheightE	SCOS	–0.31	2.49	–2.40	2.38	–0.04	3.91	0.15	3.35
	Vibrato	0.45	3.11	–2.10	2.23	0.99	3.87	0.25	3.94
ΔSFheightE%	SCOS	–0.48	19.56	–431.23	2339.63	–33.85	221.40	392.39	4090.43
	Vibrato	8.98	45.14	–122.97	201.97	18.12	62.57	111.80	1058.18
ΔSlopeE	SCOS-vibrato difference	–1.33	2.60	1.58	2.02	–1.23	4.13	–0.69	2.82
ΔSlopeE%	SCOS-vibrato difference	–25.68	58.41	22.02	25.37	–81.33	198.82	–9.33	39.85

B

	t Test for Equality of Means							
	Chanter1		Chanter2		Chanter3		Chanter4	
Slope/SPFheight	df	Sig. (2-tailed)	df	Sig. (2-tailed)	df	Sig. (2-tailed)	df	Sig. (2-tailed)
ΔSlopeE	327	**0.000**	322	**0.000**	239	**0.023**	197	0.089
ΔSlopeE%	295,216	**0.000**	17.781	**0.009**	135.524	**0.005**	197	0.101
ΔSFheightE	327	**0.015**	322	0.603	239	**0.043**	197	0.841
ΔSFheightE%	210,332	**0.016**	322	0.577	115.547	**0.022**	197	0.498

Notes: Significance level of $\alpha = 0.005$. The ending letter "E" means the "Energy" method.
* Also, the respective percentage values are demonstrated for comparison. The bold characters indicate the statistically significant differences between vocal ornament and its neighbor steady note.
† The bold characters indicate the statistically significant differences between SCOS and vibrato. Also, the respective percentage values are presented for comparison.
‡ Δ prefix: Difference between vocal ornament and steady note value.
Abbreviations: df, freedom degrees; SCOS, single cycle ornamentation structure; Sig., Sigma value.

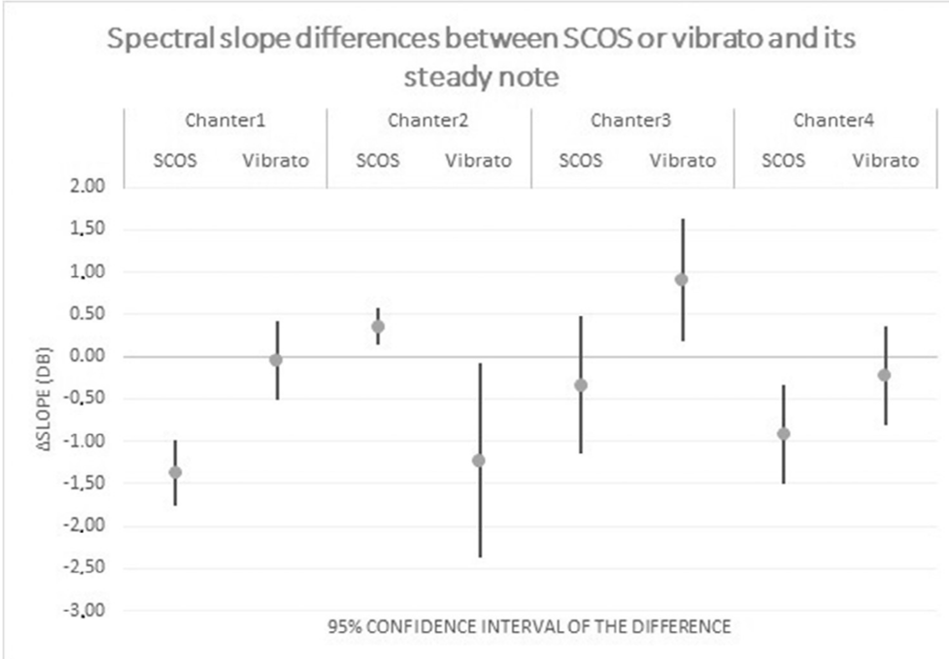

FIGURE 4. Spectral slope differences between each vocal ornament (single cycle ornamentation structure [SCOS] and vibrato) and its neighbor steady note for all four chanters. The slope difference values for SCOS tend to be lower than for vibrato values for chanters 1, 3, and 4, but for chanter 2, the corresponding values are higher for SCOS than those for vibrato.

Parameters' dependency on vocal ornament, vowel, pitch, and intensity

Initially, the role of intensity for the vocal ornament and its steady note was studied. It was found that the mean differences of intensity between the vocal ornament and the steady note did not exceed the absolute value of 0.7 dB for all four chanters 1, 2, 3, and 4: [81.6(±11.5)–80.9(±14.4) dB, 88.1(±4.4)–87.4(±4.1) dB, 91.0(±4.5)–91.3(±4.6) dB and 93.6(±4.0)–93.7(±3.7) dB, respectively]. Hence, the factor of intensity could be considered not important when comparing the parameter

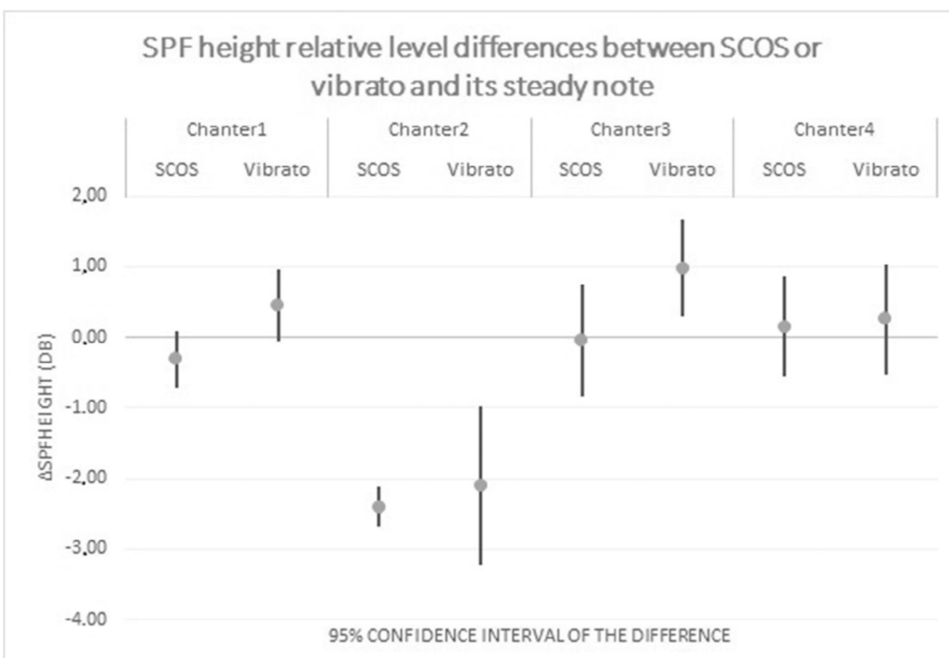

FIGURE 5. Speaker's formant height relative level differences between vocal ornament, either single cycle ornamentation structure (SCOS) or vibrato, and its steady note for all four chanters. The statistically significant differences of the mean values for chanters 1 and 3 are slightly lower in SCOS (by about 1 dB) than in vibrato values, whereas for the other two chanters, the differences are negligible due to the large variability of their values.

values between the vocal ornament of SCOS or vibrato and its neighbor steady note. However, the respective percentage values of the spectral parameters could be affected by intensity due possibly to changes in the parameters' values of the steady note (Equation 2).

By implementing a univariate analysis of the variance, we tested the effects of vocal ornament type, vowel, and pitch on the rate, the extent, and the SS and SPF height relative level differences. It was found that the extent (1) strongly depends on the vocal ornament type, for all chanters and both methods of measurements, (2) depends on the vowel (/a, e, I, o, u/), only for chanter 2 and method 2, and (3) depends on the pitch for only chanter 1 and chanter 2 and for both methods of measurements (Supplementary Appendix C2(a)). These results confirm that the extent of SCOS differs from vibrato extent, and they suggest that vowel and pitch do not affect the extent of the two vocal ornaments in a unified way for all chanters; they may depend on the personal chanting style of each one of the chanters.

The rate depends on (1) the type of vocal ornament for only one of the methods of measurements and not for all chanters, (2) the vowel, for only chanter 2 and chanter 4, when method 1 was applied, and (3) the pitch only for the cases of chanter 3 and method 2 (Supplementary Appendix C2(b)). According to these findings, the trend of differing the SCOS from vibrato in the rate does exist.

SS differences clearly are affected by (1) the vocal ornament type for three of four chanters and (2) the pitch, which is related more or less to the intensity for each chanter. Those chanters who had a better control of their voice performed higher pitch notes with small increase in intensity.[6(p39–42)] The vowel factor does not explicitly affect the SS differences for each chanter (Supplementary Appendix C2(c)).

Speaker's formant height relative level differences depend on (1) vocal ornament type, (2) vowel, and (3) pitch for most of the chanters (Supplementary Appendix C2(d)). The general trend for these mean difference values was found for SCOS to be lower than that for vibrato. Finally, the NHR and HNR were found to depend on (1) the vocal ornament type, for all chanters, (2) the pitch, for some chanters, and (3) rarely, the vowel (Supplementary Appendix C2(e)).

Formant tuning for SCOS and vibrato

The formant tuning effect was investigated through the percentage difference between the frequency of the *n*th order harmonic augmented by the extent and the first formant frequency (F1) while dividing by the half of the F0 (Supplementary Appendix C1). As the pitch range of the analyzed BC melodies varied between 106 and 344 Hz, and the average F1 was, roughly, equal to 500 Hz, only for the second (H2) or the third partial (H3) could F1 come close enough in order to observe a formant tuning effect. For the H2, this percentage difference with respect to the half of the F0 was found to vary between the extremely low value of 3.08% (large proximity between F1 and 2*F0) and the simply low value of 20.24% (Table 6). For the H3, this percentage difference in absolute values was found to be higher

TABLE 6.
Percentage Mean Differences Between the Second (H2) or Third Harmonic (H3) and the First Formant Frequency (F1) with Reference to the Half of Fundamental for All Chanters

Partial	Chanter 1			Chanter 2			Chanter 3			Chanter 4		
	Mean Difference	95% Confidence Interval of the Difference		Mean Difference	95% Confidence Interval of the Difference		Mean Difference	95% Confidence Interval of the Difference		Mean Difference	95% Confidence Interval of the Difference	
(extent)	(%)	Lower	Upper	(%)	Lower	Upper	(%)	Lower	Upper	(%)	Lower	Upper
H2 (extent1)	**−20.24**	−27.08	−13.39	−4.15	−13.52	5.23	3.08	−3.12	9.28	−7.31	−17.20	2.57
H2 (extent2)	**−23.44**	−30.30	−16.59	−5.34	−14.35	3.68	3.08	−2.94	9.10	−8.12	−18.14	1.90
H3 (extent1)	**29.83**	21.31	38.35	**31.91**	25.91	37.91	**60.92**	51.85	70.00	**32.92**	24.50	41.33
H3 (extent2)	**29.14**	20.52	37.77	**26.80**	20.74	32.86	**58.77**	49.41	68.14	**30.24**	21.82	38.66

Notes: The bold characters indicate the statistically significant percentage frequency differences from the F1, positive or negative. The absolute values are lower for the second harmonic than for the third harmonic for each chanter.

than that for the H2 for the same extent and chanter. Through a one-sample t test for the percentage differences between the H2 and F1, statistically significant differences were found only for chanter 1, whereas between H3 and F1, statistically significant differences for all chanters were found. When the difference between H2 or H3 and F1 is statistically significant, the chanters may tune F1 just above the H2 (chanter1) or just below the H3 (all chanters), whereas if the difference between H2 or H3 and F1 is not statistically significant (small difference), they may tune F1 exactly as H2 (chanters 2, 3, and 4: Table 6—bold characters).

DISCUSSION

Rate and extent

Because the extent was measured as the average value of the extent contour, namely, the SCOS was considered as a pseudo-sinusoidal with an offset equal to the distance from the steady note's baseline, this distance is twice the extent value. Hence, the frequency peak of the SCOS can be calculated as the double value of the extent one, while its average value varies from 1.96 to 2.30 ST with standard deviation of about 0.83 ST (Table 3A). This range of values is contained in the range of 1.6–2.7 ST, namely, between the trill and trillo extent values,[5] although there is an overlapping with the Indian mordent (1–2.8 ST).[53] The two methods of the rate and extent measurements produced slightly different results (roughly at most 20%) because of the way of the intonation's track estimation; however, for both methods, the extent of SCOS was systematically higher than that of vibrato and yet, regarding the rate, the differences between SCOS mean values and vibrato had the same sign for both methods (Figure 3). The rate values of SCOS, although being within the range of the vibrato rate values (3–8 Hz), showed a trend to be slightly different from those of vibrato and the sign of these differences depended on the chanter, probably because of the particular artistic style of chanting. The mean rate for chanter 4 (4.01 Hz) was notably smaller than the general average of the rate (5.11 Hz) for the SCOS, and this might mean a special way of performing that vocal ornament. However, such a conclusion needs further confirmation by involving more chanters and melodies in the analysis.

It is possible that the personal style of performing of a chanter affects the rate of the SCOS, which could be produced in a similar way as the vibrato or maybe not. Stark quotes from Seashore[1] that the same mechanism that produces the vibrato also propels vocal trills and florid vocal roulades.[11] For Seashore, "the trill is considered to be an expansion of the compass of the vibrato to the extent that two pitches are perceived rather than one; the same physiological mechanisms are present in both, and the trill rides on the impulse of the vibrato."[1] However, this comparison between trill and vibrato assumes the existence of vibrato. As in the case of chanter 2, there were many SCOS samples but only a few vibrato, and this is generally true for a number of BC chanters, we can speculate that chanters resembling chanter 2 do not produce SCOS in the same way as the vibrato. Furthermore, given that the SCOS extent values were found, for all methods and chanters, to be higher (about double values) than vibrato, this type of vocal ornament might be produced in the larynx in a different way than vibrato.

Spectral parameters for SCOS and vibrato

The percentage and nonpercentage differences of NHR and HNR showed statistically significant values for all chanters and vocal ornaments, namely, the same results. The mean values of NHR for the steady note, SCOS, and vibrato can be estimated through Equations (1) and (2); they vary in the ranges [0.0025–0.0039], [0.037–0.049], and [0.012–0.019], respectively. The corresponding ranges for HNR also can be estimated: [26.11–26.79 dB], [16.52–17.86 dB], and [21.76–22.26 dB]. The NHR values for male subjects in Fado singing varied between 0.09 and 0.12,[37] namely, more than twice the values from this study; a fact that could have the possible explanation of the different methods of measure used here.[37] On the contrary, Butte et al[36] found that SNR (similar to HNR) gets values inside the same range as in this study; for the pop singing style, the lowest value was 12.3 dB, and for the other opera, soul, country, jazz, and music theater, the values were inside the range [17.1–24.4 dB]. To interpret the higher values of the NHR and the lower values of the HNR in the case of the SCOS compared with the vibrato ones, we suggest that a less harmonically vibration of the vocal cords takes place during the SCOS performance compared with the vibrato case, a thing that indicates a rather special functioning of the larynx for the SCOS production.

The standard deviations of the percentage difference values of the SS between the SCOS and its neighbor steady note, compared with their mean values, were different from that of the nonpercentage values, whereas for the three chanters, the standard deviation was smaller in the case of percentage values (Table 5A). This observation could mean that the variability of the SS values for the steady note affects the variability of the percentage values (Equation 2). Also, the same changes were observed for the SS for vibrato values. However, these changes with respect to the variability between the percentage and no-percentage values are made in such a way that the relation between the SCOS and vibrato remains unchangeable. For example, although the mean difference between SCOS and steady note for the nonpercentage values of chanter 2 was statistically significant, that for percentage value was not statistically significant. Despite that, the relation between SCOS and vibrato for the percentage values is similar to that for nonpercentage values (Tables 5). If considering the dependency of SS on pitch as previously reported, these changes may be due to the intensity changes. In any case between SCOS and vibrato, the sign of the difference in the statistically significant differences was the same for both percentage and non-percentage values for the first three chanters. As for chanters 1 and 3, the SS differences were lower for SCOS than for vibrato, and provided that the SS depends on the voice source's operation and rather on subglottal pressure,[18] it is reasonable to suggest that SCOS is rather produced with a more "relaxed" larynx than vibrato for these BC chanters. Furthermore, both SCOS and vibrato may be performed with smaller subglottal pressure than in their steady notes. Given that the chanters for this study were selected to belong to different groups of BC styles which incorporate the most frequently appeared features in BC,[44] the above may be true not only for two of four chanters but also for many chanters not included in this study. With respect to chanter 2, the different sign of the SS difference between SCOS and vibrato, it seems that he, and some BC chanters resembling to him, may use a different vocal source mechanism to produce the SCOS and vibrato than chanters 1 and 3.

The SPF height relative level differences, percentage or nonpercentage values, between SCOS or vibrato and the steady note for chanters 1, 3, and 4 showed a large variation of the values around zero, from 7 to 39 times bigger than their mean values, respectively (Table 5A). In other words, there was not an explicit trend of mean values to be different from zero value, on average. However, for chanter 2, if taking into account that for SCOS, the nonpercentage values varied in the range [–4.78, –0.02 dB] whereas the standard deviation of the percentage values was about five times of their mean value, we can consider that, for this chanter, there is a trend for lowering the SPF level of SCOS compared with its neighbor note. Similar to the SCOS reasons, the SPF relative level in the vibrato production does not change on average compared with its associated steady note for chanters 1 and 4, but, for chanters 2 and 3, there exists a tendency for lowering the SPF level in comparison to its steady note (Table 5A—bold characters). Therefore, from the SPF height relative level measures, for the SCOS production, just one chanter appears with an explicit trend for lowering of SPF relative height level who may use a rather different way of voice functioning than others for the SCOS production (Figure 5).

The trend of the SPF relative level for lowering in SCOS compared with vibrato was clear for the nonpercentage values and chanters 1 and 3, whereas for the percentage values, that trend also exists for the first three chanters. Regardless of the small mean differences in the SPF level, this depends on the chanter and the type of vocal ornament. It has been reported that the SPF level depends on singer proficiency, vowel, fundamental frequency, vocal loudness, and phonation mode.[54] The dependence of the SF on vowel has also been found for male opera singers.[55] Furthermore, the SPF relative level can decrease if two higher formants are moved away from each other.[14] The tendency for the SCOS to be slightly lower than the SPF height relative level, compared with vibrato, is possibly related to the phonation mode as well as to some moving of the F3 and/or F4 toward the higher frequencies as someone could see in Supplementary Tables A1 and B1, other factors being unchangeable between the vocal ornament and the steady note.

The personalized statistically reliable differences for the formant frequencies and bandwidths, found for each chanter, strengthen the assumption that some results depend on the personal chanting style. The observation that almost all the ΔF1 mean differences are negative (Supplementary Table B1) indicates that the first formant frequency (F1) of the both vocal ornaments tends to be lower compared with its corresponding steady note. As the second harmonic (H2) is often lower than the F1 and the F0 remains unchangeable between the vocal ornament and its neighbor note, this means that F1 tends to get more close to H2, thus indicating the existence of a formant tuning effect during the performance of both SCOS and vibrato.

SCOS and trill or trillo

A final issue about the SCOS was whether or not it had a relation with the trill and/or trillo. Lacking identified such ornaments in BC (at least, none of the participated chanters was aware of such musical terms), it was hard to associate the SCOS with the trill and/or trillo. However, in the Western opera, there were some studies in the trill and/or trillo, in which the trill was specified as having the same rate as the vibrato rate[4] (Hak-1990), while its extent could arrive up to 2.7 ST. Instead, the trillo could have various rate values from 2 to 12.4 Hz, whereas its extent was smaller than trill (1.6 ST). Other vocal ornaments from non-Western singing styles such as the Iranian tahrir,[8,9] the Mongolian trill,[10] and the Indian mordent,[53] which resemble SCOS, were studied in the last decade. The most striking characteristic of these trills was the alternation between the modal and the falsetto registers for the two alternative notes of the trill.[8–10] The extent for these trills was measured as much as 2–7 ST for the tahrir and 3.7 ST for Mongolian trill. For the Indian mordent, the extent varied in the range of 1–2.5 ST and its time duration in the range of 110–200 milliseconds, namely, its rate varied from 5 to 10 Hz.[53] When calculating the maximum pitch in a SCOS as double the value of its extent, it is readily seen that the pitch raises by as much as 2.4 ± 0.5 ST compared with the pitch of the steady note; namely, ranging from a music interval of major second up to a minor third. Because these music intervals are met very frequently in BC, it was reasonable to assume that an SCOS was essentially a “momentary” jump to a higher successive note in the music scale. By considering the SCOS time length (0.19 ± 0.04 second, 0.17 ± 0.05 second, 0.19 ± 0.04 second, and 0.25 ± 0.04 second) together with an average tempo as much as 79.5 beats/min (70, 88, 78, and 82 beats/min), it appeared to be equivalent to an almost quarter note: $0.2/0.75 \approx 1/4$. However, the time length of maximum pitch in SCOS was less than the total SCOS time duration and could not be exactly specified. If supposing the maximum pitch of SCOS corresponds to a higher than the steady note, this note’s duration is equivalent to at least an eighth note. For the higher note not to be fused with the lower steady note of an SCOS, the pitch difference between them had to be larger than the glissando threshold that corresponds to the half of SCOS time duration. As the glissando threshold of a pitch change, for a time duration of 1 second, is about 2 ST,[13] and the SCOS extent varies between 1.9 and 2.9 ST, these higher and lower pitches tend to be not fused. However, due to small time length of the higher note, its pitch is unspecified. Hence, the SCOS does not seem to be a trill with two well-established (lower–higher) alternative notes but it has similarities with the other trills particularly in the rate and extent. From our results, it cannot be inferred that any change in register takes place during the SCOS performance; hence, the SCOS music gesture in BC seems to be a special vocal music ornament.

CONCLUSIONS

In this study, a special vocal ornament frequently performed in BC melodies was described and acoustically analyzed. This vocal ornament while somewhat resembling a one cycle period vibrato was considered as a single cycle pseudo-sinusoidal that is followed or/and preceded by a steady note; for this reason, it has been defined as the SCOS. The SCOS compared with the vibrato ornament was found to be different: (1) in the extent of which the values for SCOS were found for all four chanters to be higher almost double than the corresponding values for vibrato, (2) in the rate though with not the same sign of the difference (depending on chanter and method), (3) in the NHR that is higher for SCOS compared with vibrato and inversely in the HNR, (4) in the SS although with not the same sign of difference (depending on chanter), and (5) in the SPF relative level, which is slightly lower in SCOS, for half of chanters. Some personalized statistically reliable differences between SCOS and vibrato were found for the formant frequencies and bandwidths, and for both ornaments of SCOS and vibrato, the first formant frequency tends to decrease compared with steady note, and a formant tuning effect is observed.

Acknowledgments

The authors wish to thank the anonymous reviewers for their helpful suggestions.

SUPPLEMENTARY DATA

Supplementary data related to this article can be found online at doi:10.1016/j.jvoice.2017.10.016.

REFERENCES

1. Seashore CE. *The Vibrato*. Iowa City, IA: University of Iowa Press; 1932.
2. Shipp T, Doherty ET, Haglund S. Physiologic factors in vocal vibrato production. *J Voice*. 1990;4:300–304.
3. Horii Y. Acoustic analysis of vocal vibrato: a theoretical interpretation of data. *J Voice*. 1989;3:36–43.
4. Hakes J, Shipp T, Doherty ET. Trillo rates exhibited by professional early music singers. *J Voice*. 1990;4:305–308.
5. Hakes J, Shipp T, Doherty ET. Acoustic characteristics of vocal oscillations: Vibrato, exaggerated vibrato, trill and trillo. *J Voice*. 1988;1:326–331.
6. Sundberg J. *The Science of the Singing Voice*. DeKalb, IL: Northern Illinois University Press; 1987:76.
7. Sataloff RT. *Professional Voice: The Science and Art of Clinical Care*. New York: Raven Press; 1991.
8. Castelengo M, During J, Lamesch S. The Iranian tahrir: acoustical analysis of an ornamental vocal technique. In: Maimets-Volt, K, Parncutt, R, Marin M, Ross J, eds. *Proceedings of the 3rd Interdisciplinary Musicology Conference, CIM07. Singing*. 15-19-August, Tallinn: Estonia; 2007:1–3.
9. Biglari HJ. *Timbral and melodic characteristics of the Persian singing style of Avaz* [Master Thesis]. Stockholm, Sweden; 2012.
10. Pilliot-Loiseau Cl, Crevier-Buchman L, Rialland A, et al. Vocal registers transitions phenomena in trills used in Mongolian Long Song productions. In: Antoine Giovanni et Nathalie Henrich. *Proceedings of the 9th Pan European Voice Conference, Marseille, France*. 1(1). 2011:22–23, 2011. <halshs-00674662>.
11. Stark JA. On the role of vocal idioms in singing. *Can Univ Music Rev*. 1995;15:70–90.
12. Rapoport E. Emotional expression code in opera and lied singing. *J New Music Res*. 1996;25:109–149.
13. D'Alessandro C. The pitch of short-duration vibrato tones. *J Acoust Soc Am*. 1994;95:1617–1630.
14. Fant G. *Acoustic Theory of Speech Production*. The Hague, The Netherlands: Mouton; 1960.
15. Löfqvist A. The long-time-average spectrum as a tool in voice research. *J Phonetics.* 1986;14:471–475.
16. Cleveland T, Sundberg J, Stone R. Long-term-average-spectrum characteristics of country singers during speaking and singing. *J Voice*. 2001;15:14–60.
17. Fant G, Liljencrants J, Lin Q. A four-parameter model of glottal flow. *STL-QPSR, KTH*. 1985;4:1–13.
18. Nordenberg M, Sundberg J. Effect on LTAS of vocal loudness variation. *TMH-QPSR, KTH*. 2003;45:93–100.
19. Löfqvist A, Mandersson M. The long-time average spectrum of speech and voice analysis. *Folia Phoniatr (Basel)*. 1987;39:221–229.
20. Omori K, Kacker A, Carroll LM, et al. Singing power ratio: quantitative evaluation of singing voice quality. *J Voice*. 1996;10:228–235.
21. Thorpe C, Cala S, Chapman J, et al. Patterns of breath support in projection of the singing voice. *J Voice*. 2001;15:86–104.
22. Bele I. The speaker's formant. *J Voice*. 2006;20:555–578.
23. Sundberg J. Articulatory interpretation of the "singing formant". *J Acoust Soc Am*. 1974;55:838–844.
24. Titze IR. Acoustic interpretation of resonant voice. *J Voice*. 2001;15:519–528.
25. Kojima H, Gould WJ, Lambiase A, et al. Computer analysis of hoarseness. *Acta Otolaryngol*. 1980;89:547–554.
26. Boersma P. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. Proceedings of the Institute of Phonetic Sciences 17: 97–110. University of Amsterdam. 1993.
27. Sundberg J. Formant technique in a professional female singer. *Acustica.* 1975;32:89–96.
28. Schutte HK, Miller DG, Duijnstee M. Resonance strategies revealed in recorder tenor high notes. *Folia Phoniatr Logop*. 2005;57:292–307.
29. Neumann K, Schunda P, Hoth S, et al. The interplay between glottis and vocal tract during the male passagio. *Folia Phoniatr Logop*. 2005;57:308–327.
30. Sunberg J, Lã FMB, Brian PG. Formant tuning strategies in professional male opera singers. *J Voice*. 2013;27:278–288.
31. Prame E. Vibrato extent and intonation in professional Western lyric singing. *JASA*. 1997;102:616–621.
32. Arroabarren I, Zivanovic M, Bretos J, et al. Measurements of vibrato in lyric singers. *IEEE Trans Instrum Meas*. 2002;51:660–665.
33. Sundberg J, Gu L, Huang Q, et al. Acoustical study of classical Peking opera singing. *J Voice*. 2012;26:137–143.
34. Stone RE, Cleveland JTF, Sundberg PJ, et al. Aerodynamic and acoustical measures of speech, operatic, and Broadway vocal styles in a professional female singer. *J Voice*. 2003;17:283–297.
35. Manfredi C, Barbagallo D, Baracca G, et al. Automatic assessment of acoustic parameters of the singing voice: applications to professional Western operatic and jazz singers. *J Voice*. 2015;29:51.e1–517.e9.
36. Butte CJ, Zhang Y, Song H, et al. Perturbation and nonlinear dynamic analysis of different singing styles. *J Voice*. 2009;23:647–652.
37. Mendes AP, Rodrigues AF, Guerreiro DM. Acoustic and phonatory characterization of the Fado voice. *J Voice*. 2013;27:655.e9–655.e15.
38. Chadjidouli K. *History of Rembetiko (in Greek)*. Nefeli Editions. Athens: 1990.
39. Delviniotis SD, Kouroupetroglou G, Theodoridis S. Vibrato analysis in Byzantine music. In: *Proceedings of the Stockholm Music Acoustics Conf. SMAC*. Stockholm: KTH; 2013:242–249.
40. Chrysochoidis G, Kouroupetroglou G, Delviniotis D, et al. Formant tuning in Byzantine chant. In: *Proceedings of the Stockholm Music Acoustics Conf. SMAC*. Stockholm: KTH; 2013:217–223.
41. Chrysanthos from Madytos. *The Mega Theoreticon (in Greek)*. Trieste: Michele Weis Press; 1832. Available in digital form from ANEMI. The Digital Library of Modern Greek Studies. Available at: http://anemi.lib.uoc.gr/metadata/d/4/7/metadata-01-0001567.tkl. Accessed October 20, 2016.
42. The Divine Music Project: Prologue. Available at: http://www.stanthonysmonastery.org/music/Prologue.htm. Gregorios Stathis, Athens Feb. 2006. Accessed October 20, 2016.
43. The Divine Music Project: Liturgy of Saint John Chrysostom. Available at: http://www.stanthonysmonastery.org/music/Cherubic.htm. St. Anthony's Greek Orthodox Monastery, Florence, Arizona. #Cherubic. Accessed October 20, 2016.
44. Delviniotis D, Kouroupetroglou G. DAMASKINOS: the prototype corpus of Greek Orthodox Ecclesiastical chant voices. In: *Proceedings of Conference Crossroads|Greece as an Intercultural Pole of Musical Thought and Creativity; Thessaloniki*. 2011:1–14.
45. PRAAT: Doing Phonetics by Computer. Available at: https://www.fon.hum.uva.nl/praat/. Version: 5.4.04. Accessed January 17, 2015.
46. MATLAB. Available at: https://www.mathworks.com/. Version: R2015b, http://www.cc.uoa.gr/texniki-ypostiri3h/egkatastash-paketwn-logismikoy/matlab.html. Accessed November, 2015.
47. PRAAT manual: Sound: Filter (pre-emphasis). Paul Boersma and David Weenink, 2015.
48. Frøkjær-Jensen B, Prytz S. Registration of voice quality. *Brüel Kjær Tech Rev*. 1976;3:3–17.
49. Boersma P, Kovacic G. Spectral characteristics of the three styles of Croatian folk singing. *J Acoust Soc Am*. 2006;119:1805–1816.
50. Delviniotis SD. Acoustic characteristics of modern Greek Orthodox Church music. *J Voice*. 2013;27:656.e1–656.e12.
51. Yumoto E, Gould W, Baer T. Harmonics-to-noise ratio as an index of the degree of hoarseness. *J Acoust Soc Am*. 1982;71:1544–1550.
52. Schotz S. Speaker Classification I: Fundamentals, Features and Methods—Google books. pp. 99. 2017. Available at: https://books.google.gr/. Accessed January 13, 2017.
53. Bonjyotsna A, Bhuyan M. Analytical study of vocal vibrato and mordent of Indian popular singers. *J Voice*. 2016;30:764.e11–764.e22.
54. Sundberg J. Level and center frequency of the singer's formant. *J Voice*. 2001;15:176–186.
55. Bloothooft G, Plomp R. The sound level of the singer's formant in professional singing. *J Acoust Soc Am*. 1986;79:2028–2033.