Formant Frequency Tuning in Professional Byzantine Chanters

Georgios Chrysochoidis and Georgios Kouroupetroglou

Abstract— In this study we present the digital signal processing of formant frequency tuning in the context of the Byzantine chant voices. The DAMASKINOS prototype acoustic corpus of Byzantine Ecclesiastic professional voices was adopted for the selection of the recordings for the analysis. We have investigated recordings from 2 different professional chanters in ascending and descending musical scales of the diatonic genre, for the five vowels /a/, /e/, /i/, /o/ and /u/. The method of analysis included a semi-automatic segmentation of the audio material, extraction of the pitch and formant frequencies in PRAAT and final post-processing in MATLAB. For pitch analysis we implement a robust analysis algorithm that performs acoustic periodicity detection on the basis of an accurate autocorrelation method. Formant values were calculated with the Burg algorithm. Results show clear evidence that chanters tend to use personal formant tuning strategies throughout their vocal range.

Keywords— singing analysis, chanting processing, formant tuning, pitch.

I. Introduction

IN 1960, Gunnar Fant presented the theory of a source-filter production model for vowels, in his work "Acoustic Theory of Speech Production" [1]. According to this model, the voice source produces a harmonic series, consisting of the fundamental frequency $f0$ and a large number of harmonic frequencies, the partials [1].

Specifically, when applied to vowel production, the speech signal could be thought as the result of the source signal, produced by the glottis, and the resonator or vocal tract filter. A linear mathematical model supports this theory, which allows for relatively simple handling of calculations. The vocal tract filter can be further considered as a linear time-invariant filter for very short periods of time, making the system even more mathematically tractable. The source-filter theory of vowel production is summarized in the following equation:

$$P(f) = U(f)\ T(F)\ R(f)$$

where P(f) id the radiated sound pressure spectrum of speech, U(f) refers to the volume velocity and is used because of vocal folds act like a source of air pulses. T represents the transfer function and R denotes the radiation characteristics. For the current work we consider the terms U(F) and R(f) to be constant a different vowels are produced. Therefore, different vowels can be described as variations in the transfer function T(f) and the radiated spectrum P(f).

The resonance frequencies of the vocal tract are called formants, designated as F1, F2, F3, etc. in ascending order, and they can be displayed as spectral peaks in the frequency response of the vocal tract filter [1]. The vocal tract has four or five important formants that are used to amplify and dampen certain frequencies. The length and shape of the vocal tract determine the formant frequencies, resulting in the production of the different vowel sounds of the radiated speech signal. The lowest two formants F1 and F2 largely determine the vowel [1], while the remaining higher order formants are related to the quality of tone [2].

Adjusting the vocal tract in order to align formants with harmonics, thereby amplifying certain portions of the vocal spectrum, is known as formant tuning [3]. This intuitive act, that is highly dependent upon the vertical laryngeal positions, has been used from trained singers in the past, in cases where the singing voice should be heard across large spaces along with other sound sources, like for example music orchestras [3].

In the past, research concentrated on the relation between the quality of the voice and the formants. Later works [4], [5], revealed the existence of the singer's formant, which can be explained acoustically as a clustering of formants, F3 and F4, or F4 and F5, and even in some cases F3, F4, and F5. The singer's formant enables a singer to be heard over an orchestra, since there is little competition from the orchestra near the frequency range of the singer's formant [4]. This is a way for the singer to save some vocal effort, in other words it results in "vocal economy" [3].

Formant tuning is considered as another vocal strategy used by trained singers when trying to produce the ideal voice and economizing on vocal effort. Several works have been

This work has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: "THALES - Investing in knowledge society through the European Social Fund", under the project: "ERASITECHNIS: Processing, Analysis, Synthesis and Context-Based Retrieval for Multimedia Music-Related Data Bases of Traditional Music and Dancing Recordings", MIS 375435.

G. Chrysochoidis is with the Division of Signal Processing and Communication, Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, GR 15784 Athens, Greece (e-mail: geoxry@di.uoa.gr).

G. Kouroupetroglou is with the Division of Signal Processing and Communication, Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, GR 15784 Athens, Greece (e-mail: koupe@di.uoa.gr).

published concerning formant tuning strategies applied by classically trained Western operatic voices [6], [2], [8], [9], as well as contemporary [10] and traditional [11] ones. Recent works [8], [9] give a description of the formant tuning literature with details about the different methods and their limitations [9]. Three trends seem to be the dominant ones: a) F1 and F2 are tuned to a partial, b) F1 and F2 are not related anyhow to harmonics of the $f0$ and c) F1 and F2 are tuned just above their nearest partial in a way so that they don't coincide.

While literature on formant tuning continues to grow for other types of vocal music genres, Byzantine Ecclesiastic chant voice hasn't been studied extensively in the same context. Following our work on the existence of formant tuning in Byzantine chant [7] our current investigation tries to answer whether formant tuning strategies exist in Byzantine chant, which is examining if chanters use intentionally formant tuning to produce the desired acoustical result.

II. BYZANTINE CHANT MUSIC

Byzantine Chant Music (BCM) is a religious type of monophonic vocal performance practiced mainly in churches. Its main purpose is to serve the religious needs of the Orthodox Christian worship, providing a musical accompaniment for the ecclesiastical poetry [12], [13]. BCM follows aesthetical rules formed over the course of centuries, traditionally transferred from master performers to apprentices. It is a microtonal music since it contains intervals smaller than the conventional contemporary Western theory semitone.

The temperament used in modern Byzantine chant is based on a 72 equal divisions of the octave, called moria (plural form of morio) [13]. Compared to the 12-tone equal temperament followed by the Western theory, a Western semitone would equal to 6 moria in this 72-tone equal temperament system. Byzantine music uses the term echos to refer to a specialized type of musical mode [12], [13], denoting not only the musical scale being applied in a melody with a definite "tonic" or main note called vasi (Greek word for "base"), but also specific musical phrasing [13].

There are three musical genera in Byzantine chant: the diatonic, the chromatic and the enharmonic [12], [13]. The fundamental intervallic differences between the three genera, apart from musical aesthetics, serve different musical meanings, when considered in a specific lyric context. This leads to another classification of the three genera and its subdivisions, based on the ethos of the music genre, as Chrysanthos of Madytos, one of the main three reformers of the modern Byzantine music theory, named it [12], [14].

The musical scale used by the chanters in this work belongs to the diatonic genre shown in Figure 1.

III. MATERIALS

The measurements used in our analysis were acquired using the PRAAT software. PRAAT is a valuable software tool in the field of phonetics and voice analysis in general [18]. It is a flexible tool, which provides functionality that could handle most of the tasks needed in this work. It handles all the widely used audio file formats by importing the audio files as objects on which various operations can be applied. Apart from sound files, PRAAT can create and make use of tagging files in the form of layers called tiers. Tiers consist of boundaries and intervals between them. Labels can be added at these intervals making it easy to tag sound recordings. Several tiers can form a TextGrid object. To summarize the main PRAAT features we used in our analysis: a. it handles large audio files, b. it extracts measurements of the vocal parameters using its built-in functions, c. it uses tagging layers for audio file annotation, and d. it uses a scripting language for automating processing. Indeed, PRAAT can manipulate, edit and analyze long stereo audio files. Annotation of the files is done with tagging layers, using boundaries to mark time exact points in the recording, and the intervals in-between them to insert the metadata. One of the most powerful features of PRAAT is inevitably its scripting language. This language includes variables, loops, jumps, formulas, procedures, arrays, etc. which provide the flexibility to implement complex algorithms in combination with the ready-to-use analysis commands found in PRAAT.

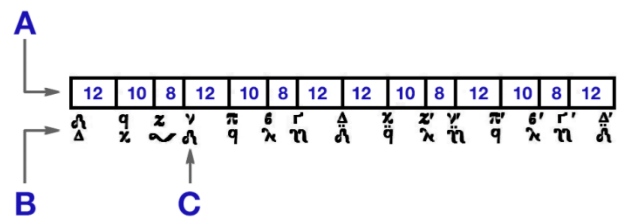


Figure 1. Diatonic genre scale used in Byzantine chant. A: Scale intervals measured in moria (72-tone equal temperament). B: Symbols used for the scale degrees. C: Scale's tonic.

The functions that were mostly used in our measurements are the ones for pitch and formant analysis. For pitch analysis PRAAT uses an algorithm that performs an acoustic periodicity detection on the basis of an accurate autocorrelation method [19].

Formant values are calculated with the algorithm by Burg [20], [21]. PRAAT is a commonly used speech analysis tool and its accuracy has been thoroughly tested in several related papers [22], [23], [24]. Furthermore, PRAAT measurements are used in many investigations as the baseline for accuracy comparisons [25], [26], [27], [28].

We started processing the audio files by using a series of PRAAT scripts that analyzed intensity and pitch, in order to annotate the audio file, by marking the basic musical units, which correspond to the scale's degrees [17]. Our final purpose was to extract separate audio files for each scale's degree. This was a three-step process: 1. voiced – unvoiced parts of the audio file were labeled using a tag layer, 2. for each voiced part boundaries were placed at pitch transitions between the scale's degrees, and 3. intervals in-between boundaries were labeled according the audio segment's average pitch. Before moving forward to the next step we were able to fine-adjust the placement of the boundaries. The transition zones between notes have been excluded. Each audio segment between boundaries was extracted to a separate audio file with average time duration of 1 sec.

Next, we analyzed each of the extracted files, using PRAAT's readily available functions, in order to acquire the actual measurements we needed. Three vocal parameters were measured every 10 msec: pitch value for f0, frequency values for the formants F1 and F2. Other measurements, like sound intensity level, formant bandwidths, formant levels and partials levels were also extracted, although not used in the current work. All data were stored in tab-delimited text files.

Final processing was done in Matlab. First the data files were imported, the mean values were calculated for pitch and formant frequencies and the corresponding graphs were plotted.

Our aim was to be able to compare both ascending and descending scales for each vowel in order to be able to determine if formant tuning was used intentionally by the chanters.

IV. Results

Displaying and analyzing the actual measurements for each chanter was the next step in our investigation. This involves plotting the frequency tracks of the two lowest formants F1 and F2, along with the partials h2-h8, where hn = n · f0 and f0 is the fundamental frequency of the vowel. Figure 2 and 3 present these measurements for each chanter respectively. Both axes represent frequency values in semitones, measured from D2 (73 Hz).

In order to be able to easily compare ascending and descending scales, the descending part was plotted using vertical mirroring as to overlay over the ascending part.

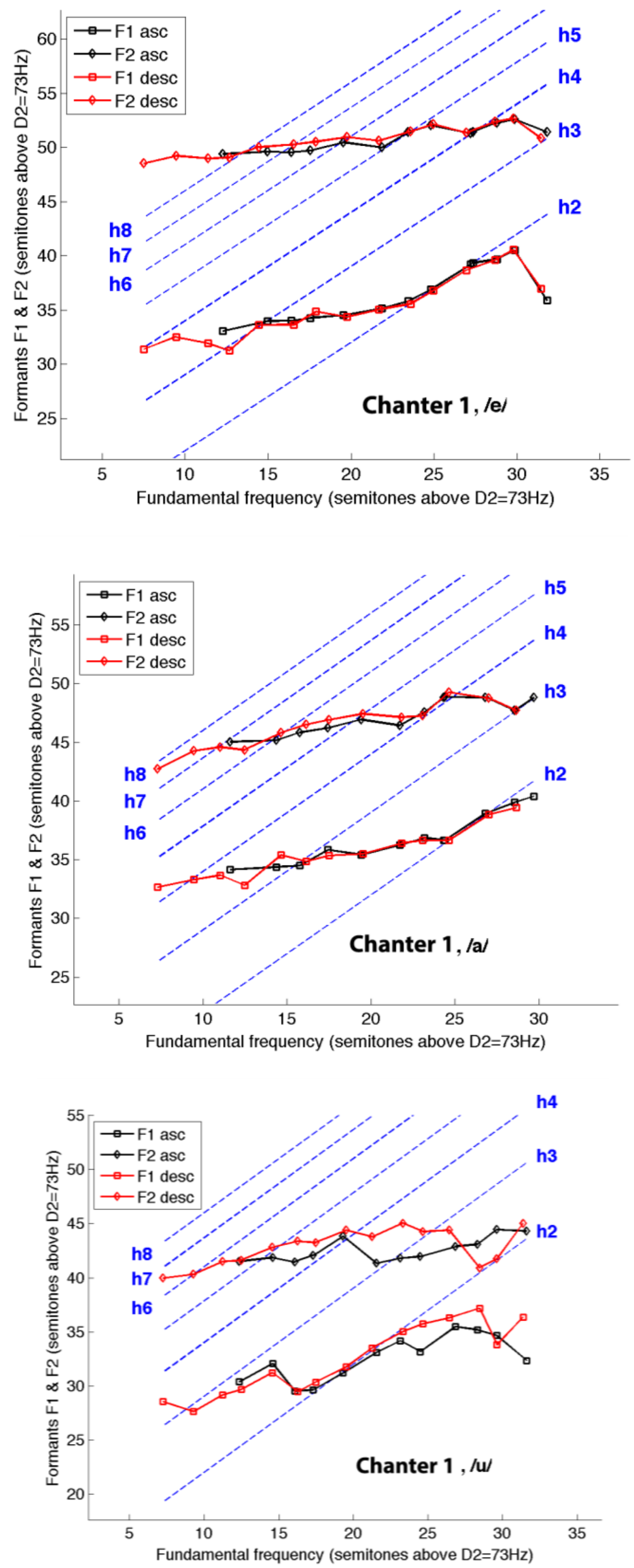
The pitch range that draws our attention is the one referred to as the passaggio, which is D4-G4 for tenors and B3-E4 for baritones [6], [9], [29]. It has been found that formant tuning is mostly observed in and above this range [9]. This can be easily seen in our results. Most chanters tend to tune their F1 and F2 to partials at this pitch range, although there were cases of formant tuning at lower scale degrees.

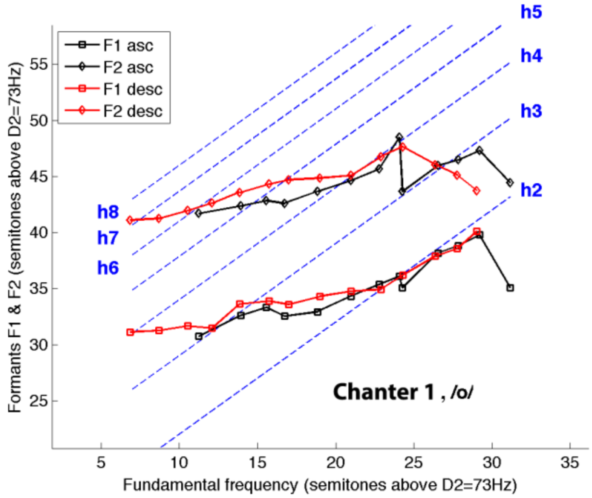
Considering an accuracy margin of about 20 Hz for the tuning of the formant frequencies [32], [9], as well as the maximum distance criterion of 50 Hz between the formant and its nearest harmonic, used in similar works [9], a 2 semitone approximate distance between the formant F1 and the partial h2, in the frequency range of E4, could still be counted as tuning [9]. This could easily explain the distance between the F1 and h2, found in the results for chanter 1, since it can be considered inside the formant tuning tolerance limits.

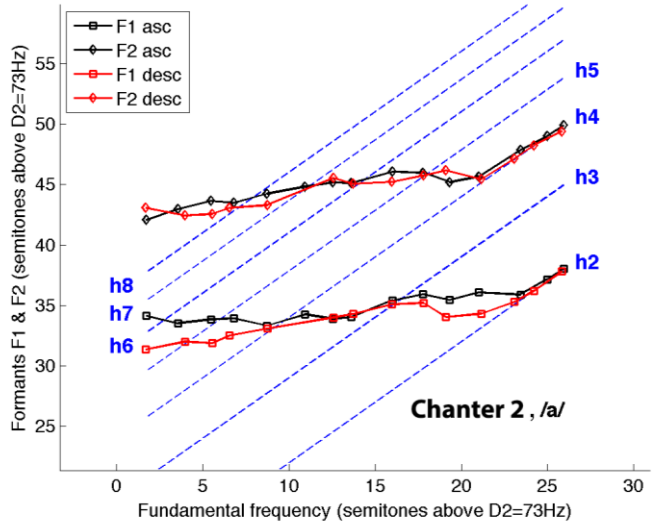
The distance in semitones between each formant F1, F2 and their closest partial for all vowels, is displayed in Fig. 2. The data presented in the graphs are for the ascending and descending diatonic scales, for both chanters.

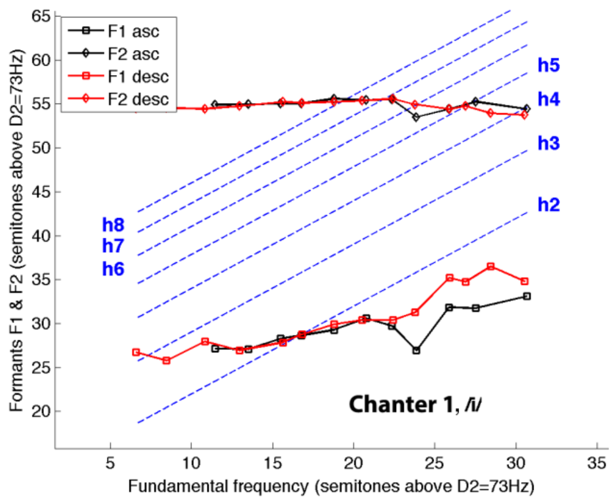
Results for chanter 1 (Fig. 2) show a clear tendency of the F1 and F2 formants to follow the slope of the h2 and h4 respectively, inside the passaggio area, although sometimes h3 and h4 are also used for the tuning of f1 and f2. Formant tuning is evident in four of the total five vowels.

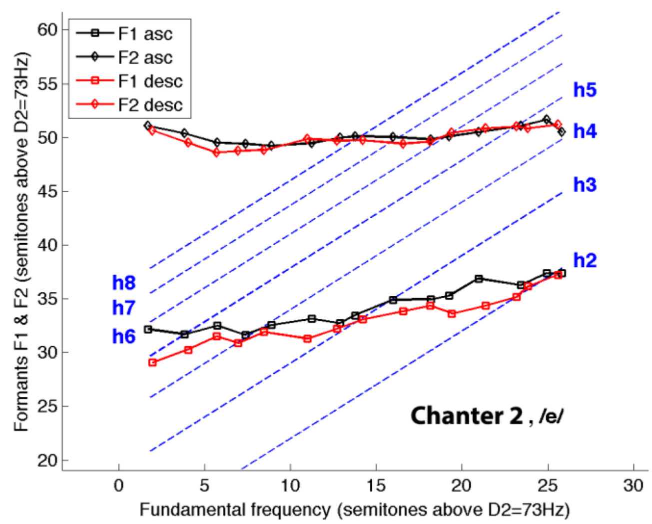
Formants F1 and F2 of chanter 2 (Fig. 3) show signs of formant tuning especially at the passaggio range and above. The second chanter represents a typical example of the formant tuning phenomenon since f1 is in most cases aligned with the h2 harmonic in all 5 vowels.











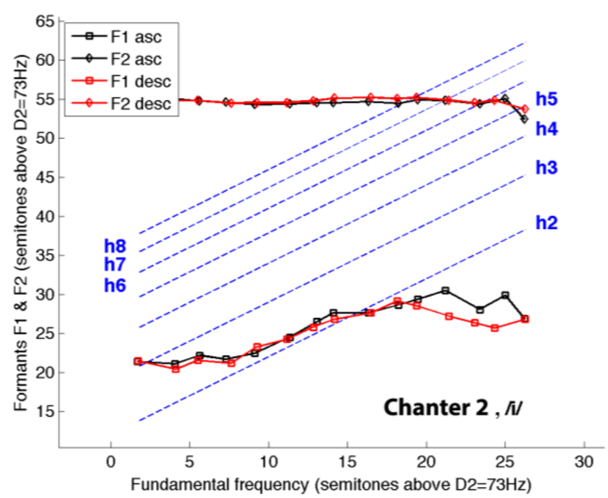


Figure 2. Frequencies, in semitones from D2 (73Hz), of the formants F1 and F2 for the vowels /a/, /e/, /o/, /u/ and /i/ for the first chanter. Results are presented for each chanter for the ascending (black line) and descending (red line) diatonic scales for all 5 vowels. Harmonics h2-h8, where $hn = n \cdot f0$ and $f0$ is the fundamental frequency of the vowel, are displayed by the diagonal blue lines. Descending scale measurements are presented using vertical mirroring in order to easily compare results with the ones from the ascending scale.

V. DISCUSSION

The reader can easily conclude that the use of formant tuning for both chanter in all five vowels is fully intentional. In all cases the two lines for ascending and descending scales follow each other very closely.

Formant tuning strategy was most apparent in cases where the F1 and F2 remained relatively constant throughout the scale, before breaching a breaking point near the beginning of the passaggio region.

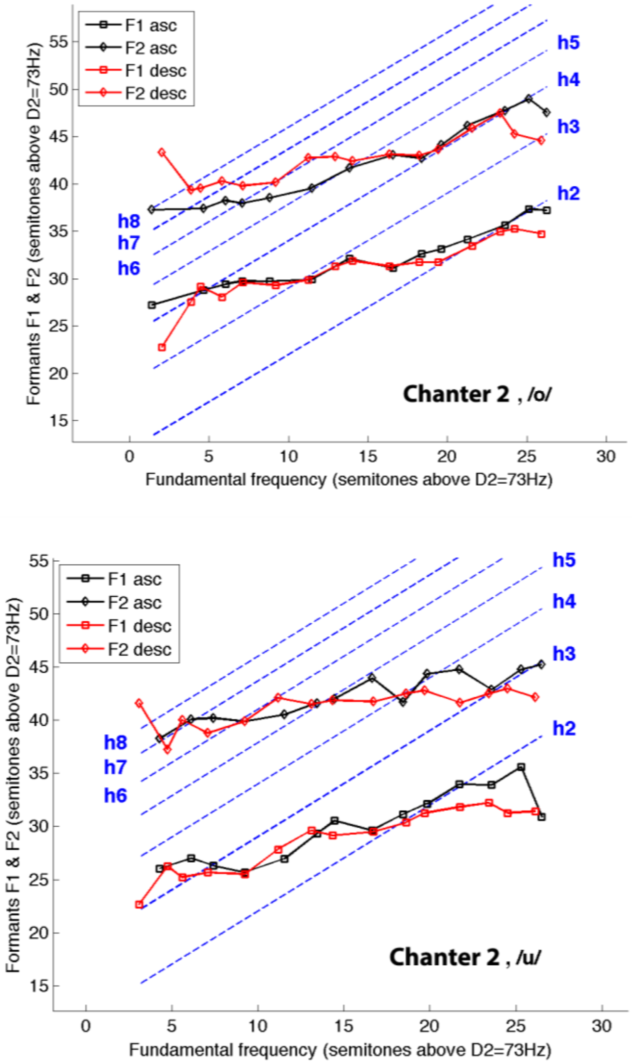


Figure 3. Frequencies, in semitones from D2 (73Hz), of the formants F1 and F2 for the vowels /a/, /e/, /o/, /u/ and /i/ for the second chanter. Results are presented for each chanter for the ascending (black line) and descending (red line) diatonic scales for all 5 vowels. Harmonics h2-h8, where h$n = n \cdot f0$ and $f0$ is the fundamental frequency of the vowel, are displayed by the diagonal blue lines. Descending scale measurements are presented using vertical mirroring in order to easily compare results with the ones from the ascending scale.

Regarding the question as to whether there is a common tuning strategy followed by most chanters, the answer is not obvious. Although in many cases we observed similarities between the chanters for F1 and F2, we tend to believe that each chanter follows his own personal strategy to achieve the aesthetic result he desires.

In order to be able to give a safe answer whether Byzantine chanters use common strategies in formant tuning a larger sample of the corpus must be analyzed. This analysis will be part of our future work.

Going through the examples of formant tuning found in our results, formant F1 coincided, in most cases, with the partial h2 and h3, while formant F2 was tuned either to h3 or h4, in all cases.

VI. FUTURE WORK

Future work will mainly cover extracting and displaying more data from the DAMASKINOS corpus. For a start, this would include retrieving data from all the chanters of the corpus. Analyzing data from other vowels as well, would give us a clear picture as to whether chanters show a preference in formant tuning for specific vowels.

Another important and interesting part of a future investigation would be the comparison of formant tuning strategies followed by BCM performers who have also been trained as professional opera singers and vice versa. The design of the DAMASKINOS corpus has included singers that meet the above conditions.

Finally, searching for a possible relation between formant tuning and different genera in BCM could also lead to important facts, both about the modern practice of Byzantine chant performance and also about the intentionality factor stated before.

REFERENCES

[1] G. Fant, *Acoustic Theory of Speech Production*. Walter de Gruyter, 1970.

[2] G. Carlsson and J. Sundberg, “Formant frequency tuning in singing,” *J. of Voice*, vol. 6, no. 3, pp. 256–260, 1992.

[3] J.A. Stark, *Bel canto: a history of vocal pedagogy*. University of Toronto Press, 2003.

[4] J. Sundberg, “The acoustics of the singing voice,” *Sci. Am.*, vol. 236, no. 3, pp. 82–91, 1977.

[5] J. Sundberg, *The Science of the Singing Voice*. Northern Illinois Univ, 1987.

[6] D.G. Miller and H.K. Schutte, “Formant tuning in a professional baritone,” *J. of Voice*, vol. 4, no. 3, pp. 231–237, 1990.

[7] G. Chrysochoidis, G. Kouroupetroglou, D. Delviniotis, S. Theodoridis: “Formant Tuning in Byzantine Chant” Proceedings of the Int. Conference Sound and Music Computer, pp. 217-223, Stockholm, Sweden, July 30-August 3, 2013.

[8] J. Sundberg, F.M.B. Lã, and B.P. Gill, “Professional male singers’ formant tuning strategies for the vowel /a/,” *Logoped Phoniatr Vocol*, vol. 36, no. 4, pp. 156–167, 2011.

[9] J. Sundberg, F.M.B. Lã, and B.P. Gill, “Formant Tuning Strategies in Professional Male Opera Singers,” *J. of Voice*, Dec. 2012.

[10] M.E. Bestebreurtje and H.K. Schutte, “Resonance strategies for the belting style: Results of a single female subject study,” *J. of Voice*, vol. 14, no. 2, pp. 194–204, 2000.

[11] N. Henrich, M. Kiek, J. Smith, and J. Wolfe, “Resonance strategies used in Bulgarian women’s singing style: A pilot study,” *Logopedics Phonatrics Vocology*, vol. 32, no. 4, pp. 171–177, 2007.

[12] D.G. Panagiotopoulos, *Theoria kai Praxis tis Byzantinis Ecclesiastikis Mousikis*, 6th ed. Athens: Sotir, 1997.

[13] D. Delviniotis, G. Kouroupetroglou, and S. Theodoridis, "Acoustic analysis of musical intervals in modern Byzantine Chant scales," *J. of the Acoustical Society of America*, vol. 124, no. 4, pp. EL262–EL269, 2008.

[14] M. Chrysanthos and P.G. Pelopides, *Theoretikon mega tes mousikes*. En Tergeste : ek tes typographias Michael Vais (Michele Weis), 1832.

[15] D. Delviniotis and G. Kouroupetroglou, "DAMASKINOS: The Prototype Corpus of Greek Orthodox Ecclesiastical Chant Voices", in *Proc. Int. Conf. Crossroads | Greece as an Intercultural Pole of Musical Thought and Creativity*, Thessaloniki, Greece, pp.1-14, 2011.

[16] G. Kouroupetroglou, D. Delviniotis, and G. Chrysochoidis, "DAMASKINOS: The Model Tagged Acoustic Corpus of Byzantine Ecclesiastic Chant Voices," *in Proc. ACOUSTICS Conf.*, Heraclion, Greece, 2006, pp. 68–76.

[17] G. Chrysochoidis, D. Delviniotis, and G. Kouroupetroglou, "A semi-automated tagging methodology for Orthodox Ecclesiastic Chant Acoustic corpora," in *Proceedings of the Int. Conf. Sound and Music Computing*, Lefkada, Greece, 2007, pp. 126–133.

[18] P. Boersma, "Praat, a system for doing phonetics by computer," *Glot International*, vol. 5, no. 9/10, pp. 341–345, 2002.

[19] P. Boersma, "Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound," in *IFA Proceedings 17*, 1993, pp. 97–110.

[20] N. Andersen, "On the calculation of filter coefficients for maximum entropy spectral analysis," *Geophysics*, vol. 39, no. 1, pp. 69–72, Feb. 1974.

[21] D. G. Childers, *Modern spectrum analysis*. IEEE Press : sole worldwide distributor (exclusive of IEEE), Wiley, pp. 252-255, 1978.

[22] P. Boersma and G. Kovacic, "Spectral characteristics of three styles of Croatian folk singing," *The Journal of the Acoustical Society of America*, vol. 119, no. 3, p. 1805, 2006.

[23] P. Escudero, P. Boersma, A. S. Rauber, and R. A. H. Bion, "A cross-dialect acoustic description of vowels: Brazilian and European Portuguese," *The Journal of the Acoustical Society of America*, vol. 126, no. 3, p. 1379, 2009.

[24] T. Wempe and P. Boersma, "The interactive design of an F0-related spectral analyser," in *Proc. 15th ICPhS*, 2003, pp. 343–346.

[25] I. Jemaa, O. Rekhis, K. Ouni, and Y. Laprie, "An evaluation of formant tracking methods on an Arabic database," in *10th Annual Conference of the International Speech Communication Association-INTERSPEECH 2009*, 2009.

[26] H. Boril and P. Pollák, "Direct time domain fundamental frequency estimation of speech in noisy conditions," in *Proceedings of EUSIPCO 2004 (European Signal Processing Conference)*, 2004, vol. 1, pp. 1003–1006.

[27] D. G. Silva, L. C. Oliveira, and M. Andrea, "Jitter Estimation Algorithms for Detection of Pathological Voices," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, no. 1, p. 567875, 2009.

[28] K. Bunton and B.H. Story, "A test of formant frequency analyses with simulated child-like vowels", Presented at the 161st Acoustical Society Meeting, 129(4), pt. 2 of 2, 2011.

[29] D. Delviniotis, "*Byzantine chant analysis by using signal processing techniques*," PhD Theses, National and Kapodistrian University of Athens, Department of Informatics and Telecommunications, 2002.

[30] L. Dmitriev and A. Kiselev, "Relationship between the Formant Structure of Different Types of Singing Voices and the Dimensions of Supraglottic Cavities," *Folia Phoniatrica et Logopaedica*, vol. 31, no. 4, pp. 238–241, 1979.

[31] T.F. Cleveland, "Acoustic properties of voice timbre types and their influence on voice classification," *J. of the Acoustical Society of America*, vol. 61, p. 1622, 1977.

[32] N. Henrich, J. Smith, and J. Wolfe, "Vocal tract resonances in singing: Strategies used by sopranos, altos, tenors, and baritones," *J. of the Acoustical Society of America*, vol. 129, p. 1024, 2011.