Formant Tuning in Byzantine Chanting

Georgios Kouroupetroglou and Georgios Chrysochoidis

National and Kapodistrian University of Athens,
Department of Informatics and Telecommunications, Greece
{koupe, geoxry}@di.uoa.gr

Abstract. Adjusting the vocal tract during singing in order to align formants (i.e. the resonance frequencies of the vocal tract) with harmonics is known as formant tuning. 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. Formant tuning is considered as another vocal strategy used by trained singers when trying to produce the ideal voice and economizing on vocal effort. While literature on formant tuning continues to grow for other types of vocal music genres [2, 3], Byzantine Ecclesiastic chant voices haven’t been studied in the same context. Current work tries to fill this gap by providing an initial approach, along the path followed, and presents the first results of this analysis. We present an investigation of formant tuning in the context of the Byzantine Ecclesiastic chant voice. The recordings selected for the analysis are part of the DAMASKINOS prototype acoustic corpus of Byzantine Ecclesiastic voices. More specifically, we analyzed recordings from ten different professional chanters in ascending musical scales of the diatonic genre, for the /a/ vowel. The method of analysis included a semi-automatic segmentation of the audio material, extraction of the measurements in PRAAT and the final post-processing in MATLAB. Results show clear evidence of formant tuning in at least 60% of the chanters, proving that the technique is in use by the modern Byzantine chanting professional performers.

Περίληψη. Το φαινόμενο της προσαρμογής της φωνητικής οδού κατά τρόπο τέτοιο ώστε να ευθυγραμμίζονται οι φωνοσυντονισμοί, δηλαδή οι συχνότητες συντονισμού της φωνητικής οδού με τις αρμονικές συχνότητες του σήματος πηγής της φωνής, είναι γνωστό ως εναρμόνιση φωνοσυντονισμών (formant tuning). Η διαισθητική αυτή λειτουργία, η οποία σε μεγάλο βαθμό εξαρτάται από την κατακόρυφη κίνηση του λάρυγγα, χρησιμοποιείται από επαγγελματίες τραγουδιστές, εκτελεστές και ψάλτες από παλιά, ειδικώς σε περιπτώσεις όπου το φωνητικό σήμα πρέπει να ακουστεί μέσα σε μεγάλους χώρους (όπερες, ναούς κλπ.) ιδίως σε συνδυασμό με άλλες ηχητικές πηγές (ορχήστρες, χορωδίες κλπ.). Η εναρμόνιση φωνοσυντονισμών θεωρείται μία μέθοδος φωνητικής στρατηγικής που χρησιμοποιείται από εκπαιδευμένους φωνητικούς εκτελεστές όταν προσπαθούν να επιτύχουν ένα ιδανικό αισθητικό αποτέλεσμα ενώ ταυτόχρονα να εξοικονομήσουν ενέργεια ελέγχοντας την φωνητική προσπάθεια που καταβάλλουν. Παρόλο του η σχετική βιβλιογραφία συνεχώς εμπλουτίζεται με δουλειές που αφορούν φωνητικά είδη που συναντώνται σε διάφορες μουσικές παραδόσεις ανά τον κόσμο, το φαινόμενο δεν έχει μελετηθεί επαρκώς στο πλαίσιο της ψαλτικής τέχνης. Η παρούσα εργασία προσπαθεί να καλύψει ένα μέρος αυτού του κενού, επιχειρώντας μία πρώτη προσέγγιση και παρουσιάζει τα αποτελέσματα. Οι ηχογραφήσεις που χρησιμοποιήθηκαν για τις αναλύσεις αποτελούν μέρους του πρότυπου σώματος ηχογραφήσεων Βυζαντινής Εκκλησιαστική Φωνής «Δαμασκηνός». Συγκεκριμένα, αναλύθηκαν ηχογραφήσεις από 10 διαφορετικούς ψάλτες σε ανιούσες και κατιούσες μουσικές κλίμακες του διατονικού γένους για το φωνήεν ‘a’. Η μέθοδος περιλάμβανε μία ημιαυτόματη μέθοδο κατάτμησης του ηχητικού υλικού, εξαγωγή των απαραίτητων μετρήσεων χρησιμοποιώντας το λογισμικό Praat, ενώ η τελική επεξεργασία των μετρήσεων έγινε σε περιβάλλον Matlab. Τα αποτελέσματα δείχνουν ότι το φαινόμενο είναι υπαρκτό σε τουλάχιστον 60% των ψαλτών, αποδεικνύοντας κατ’ αυτό τον τρόπο ότι η τεχνική χρησιμοποιείται στη σύγχρονη πράξη της ψαλτικής τέχνης.

1. INTRODUCTION

The production of voice is based on the coordination of three factors: breathing, phonation and resonance. Breathing air out of the lungs produces the power supply for the voice. This airflow from the lungs makes the vocal folds (or vocal chords) in the larynx vibrate to make the basic sound of the voice (voice source sound). This process is called phonation. The source sound travels from the larynx through the throat, mouth and nose before it is transmitted through the air. During this process the source sound is being transformed according to the characteristics of the *vocal tract* system. This transformation can be thought as a filtering process that receives the source sound and alters it by amplifying certain parts of its frequency related content. The frequencies that are amplified are known as *resonances* (*resonance frequencies*) of the vocal tract. Modifying the vocal tract and thus modifying the resonances generates the different vowels. The vowels combined with a noise production mechanism for generating consonant sounds forms the syllables used in speech and singing.

When we breathe in and out without producing a voiced sound, the vocal folds in the larynx are open to allow the air to pass to and from the lungs easily. When we intend to speak, the brain sends impulses as signals to the muscles of the larynx to close the vocal folds, thus creating tension, which is analogous to a violin string being tuned using the peg to alter the tension. When the air coming up from the lungs encounters the closed vocal folds. As a result pressure starts building up until the flow of the air overcomes the resistance of the vocal folds and sets them into a pattern of rapid vibration. That is, the vocal folds open and close repeatedly. Typical values for speech are around 190-220 times per second (Hz) for women and 100 - 140 times per second (Hz) for men [1]. This rapid vibration of the vocal folds produces sound waves, which are the basic tones of our voices. The vocal folds are therefore the source of the human voice.

In 1960, Gunnar Fant presented the theory of a source-filter production model for vowels, in his work "Acoustic Theory of Speech Production" [2]. According to this model, the voice source produces a harmonic series, consisting of the fundamental frequency f0 and a large number of *harmonic frequencies* (*harmonics*), the *partials*.

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 **Figure 1**.

The *pitch* or *fundamental frequency* f0 of the voice largely determines how the human ear perceives a sound, whether it sounds "low" (bass sound) or "high" (treble sound). 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 [2]. The vocal tract has four or five important formants that are used to amplify certain frequencies **Figure 2**. 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 [2], while the remaining higher order formants are related to the quality of tone [3].

Adjusting the vocal tract in order to align formants with harmonics, thereby amplifying certain portions of the vocal spectrum, is known as *formant tuning* [4]. 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 [4].

In the past, research concentrated on the relation between the quality of the voice and the formants. Later works [5, 6], 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 [5]. This is a way for the singer to save some vocal effort, in other words, it results in "vocal economy" [4].

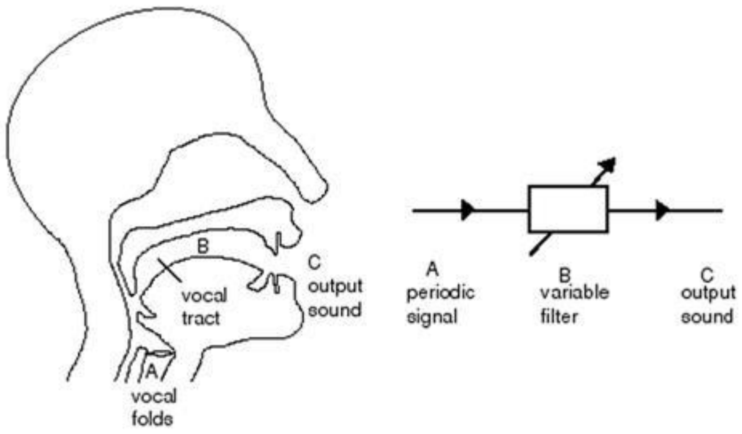

Figure 1: Source-filter production model for voice production. The speech signal could be thought as the result of the source signal, produced by the glottis, and the resonator or vocal tract filter.

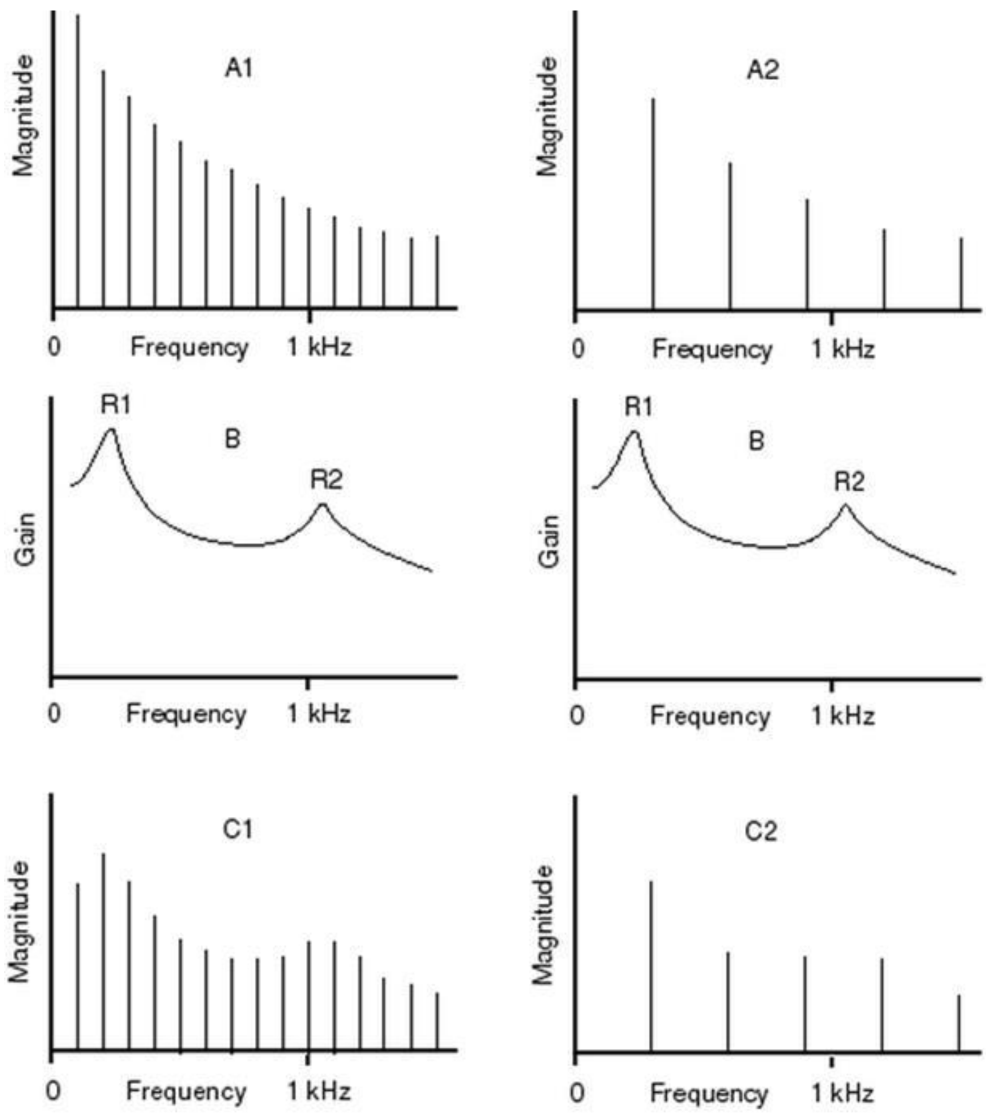

Figure 2: Two different voice source signals A1 and A2. Harmonics are shown as vertical lines. The same vocal tract filter B has been applied to A1 and A2. The result of the filtering process can be seen in C1 and C2.

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 published concerning formant tuning strategies applied by classically trained Western operatic voices [7, 8, 9, 10], as well as contemporary [11] and traditional [12] ones. Recent works [9, 10] give a description of the formant tuning literature with details about the different methods and their limitations [10]. 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, related studies for Byzantine Ecclesiastic chant voice are still at an early stage. Current work tries to fill this gap by providing an initial approach, along the path followed by other recent investigations on this subject, and presents the first results of this preliminary analysis.

2. DAMASKINOS CORPUS

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 [13, 14]. 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 materials selected to be analyzed in the present work are part of the DAMASKINOS prototype acoustic corpus of Byzantine Ecclesiastic voice [16, 17]. The DAMASKINOS corpus was designed and developed in the Department of Informatics and Telecommunications, University of Athens. Our aim was to create a standard, tagged corpus of BCM, from a statistically representative sample of modern performers. The entire corpus consists of the recordings of twenty chanters carefully selected from entire Greek territory, recorded strictly under the same controlled conditions of digital recording, with a second channel for the electroglottograph signal (EGG), as well as other measurements like glottal flow and subglottal pressure. The musical performance of each subject followed a structured protocol and included: a) specific music sections of all the musical genera from the entire repertoire of Byzantine chant (papadika, stehera, eirmolologika) and b) musical exercises in all the musical scales per genre. All the sections were recorded, except from being chanted, at their spoken and musical recitation as well, in normal intensity of voice and in the basic octave of C3 [16]. The contents of the corpus were tagged using multiple layers of metadata [18]. These data mainly consist of the lyrics syllables, the musical notation symbols of the Byzantine music score and the musical intervals. Each of these types of data forms a tagging layer on its own **Figure 3**. A couple of extra layers are used for parts of the recordings where the chanter overlooked the written score, by means of musical expression (intervals, score) [18].

In this work we used a subset of the DAMASKINOS corpus, which consists of ascending musical scales for the vowel /a/, performed by a total of 10 chanters. During recordings, the chanters were asked to perform the intervals of the diatonic scale for the specific vowel. Chanters were free to choose the performed pitch span that suits best to their vocal range.

3. METHOD OF ANALYSIS

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 [19]. 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 4**.

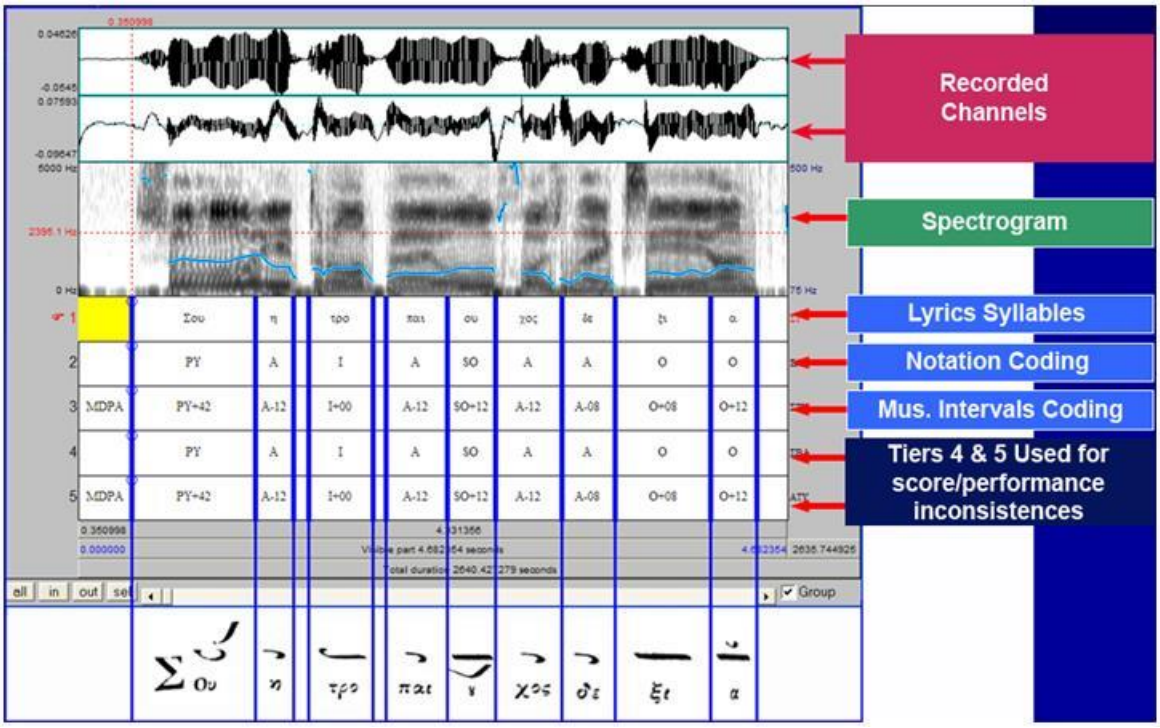

Figure 3: Short musical phrase, showing all tagging tiers.

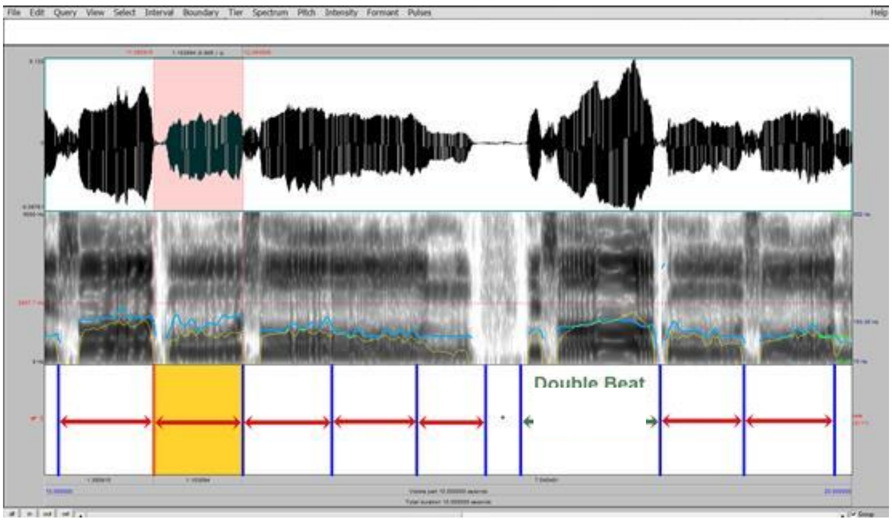

Figure 4: Analysis in PRAAT.

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 [20]. Formant values are calculated by the algorithm by Burg [21, 22]. PRAAT is a commonly used speech analysis tool and its accuracy has been thoroughly tested in several related papers [23, 24, 25]. Furthermore, PRAAT measurements are used in many investigations as the baseline for accuracy comparisons [26, 27,28, 29].

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 [18]. 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.

4. RESULTS

Displaying the actual measurements for each chanter is 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 $hn = n \cdot f0$ and f0 is the fundamental frequency of the vowel. In **Figure 5** we present these measurements for each chanter. Both axes represent frequency values in semitones, measured from D2 (73 Hz).

Results for chanter 1 show a clear tendency of the F1 and F2 formants to follow the slope of the h2 and h4 respectively, above the C4 note. This is clearly depicted by a simultaneous characteristic bend in the track of both F1 and F2 around the pitch of C4.

Formants F1 and F2 of chanter 2 follow a similar pattern. The simultaneous bend at about the same pitch, as the one noticed for chanter 1, is found in F1 and F2 of chanter 2, although F2 jumps over to the h3 partial at the scale's peak note.

Chanters 3 and 4 are two more complicated cases. While both of them show signs of formant tuning, it is not clear if this is done intentionally. Specifically, chanter 3 seems to tune F1 to h3, then to h2, while getting higher in the pitch range between C3 and D#4. On the other hand, chanter 4 shows a quite similar F1 track to the one described for chanter 3. Indeed F1 tunes initially to h3 from C3 to F3, then tunes down to h2 from A3 to D4. In both 3 and 4 chanters F2 tunes to h3 at the highest part of the scale, namely between E4 and A4.

Formants F1 and F2 for the chanters 5 and 6 seem to be unrelated to the partials. Despite the fact that for some tones they seem tuned on the harmonics, the overall image of the two formants tracks, leads us to believe that this was a coincidence and not the result of an intentional tuning effort. An exception seems to occur at the top 3-4 notes where F2 is tuned to h3 or h4.

Chanter 7 tunes its F1 to the h2 partial starting approximately at D4 and keeps it tuned up until the ultimate note of G#4. Formant F2 for the same chanter tunes to h4 at about C4 then jumps to h3 at F4.

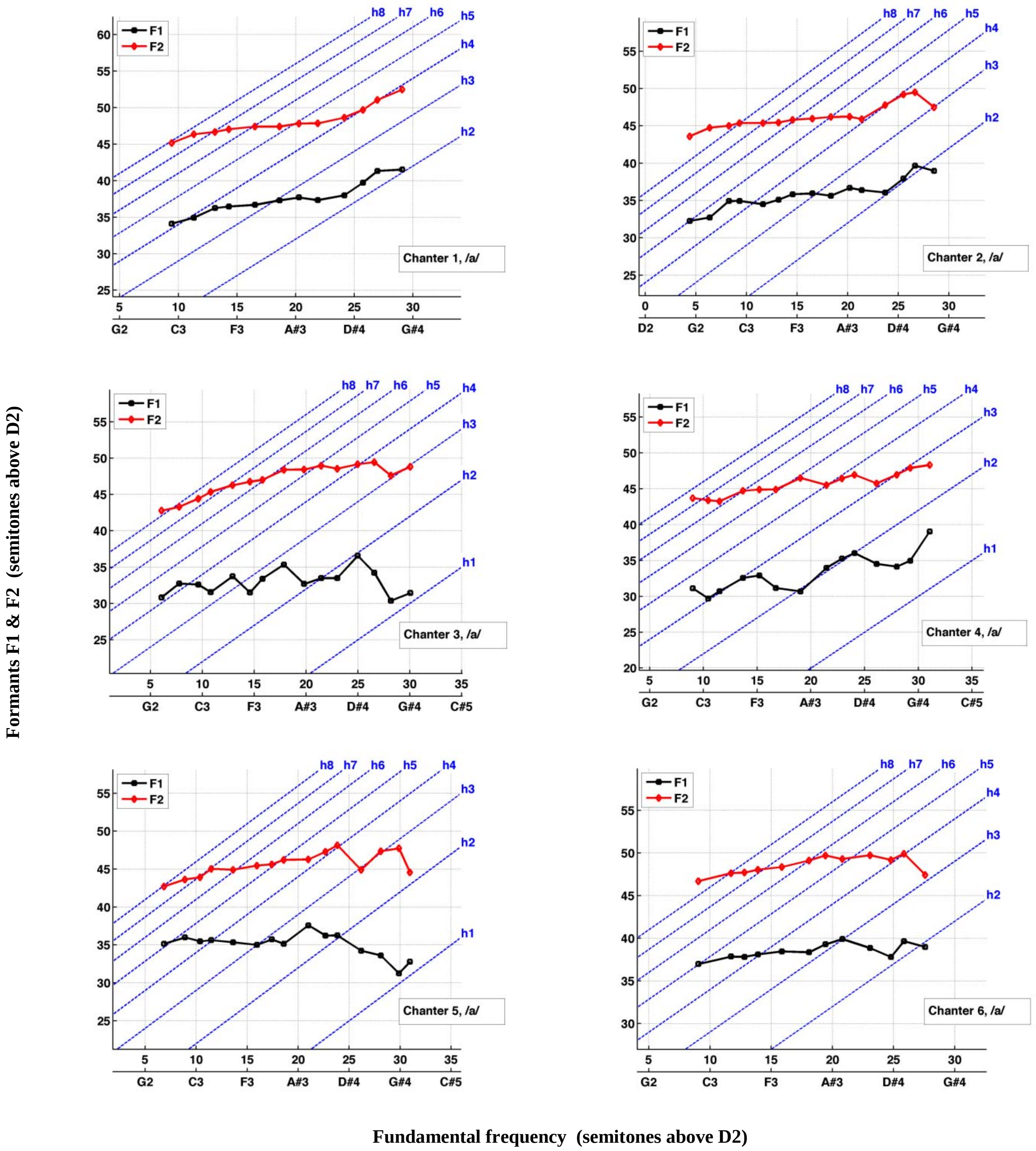

Figure 5: Frequencies, in semitones from D2 (73Hz), of the formants F1 and F2 for the vowel /a/. Results are presented for each chanter for the ascending diatonic scale. Harmonics h2-h8, where $hn = n \cdot f0$ and f0 is the fundamental frequency of the vowel, are displayed by the diagonal blue lines.

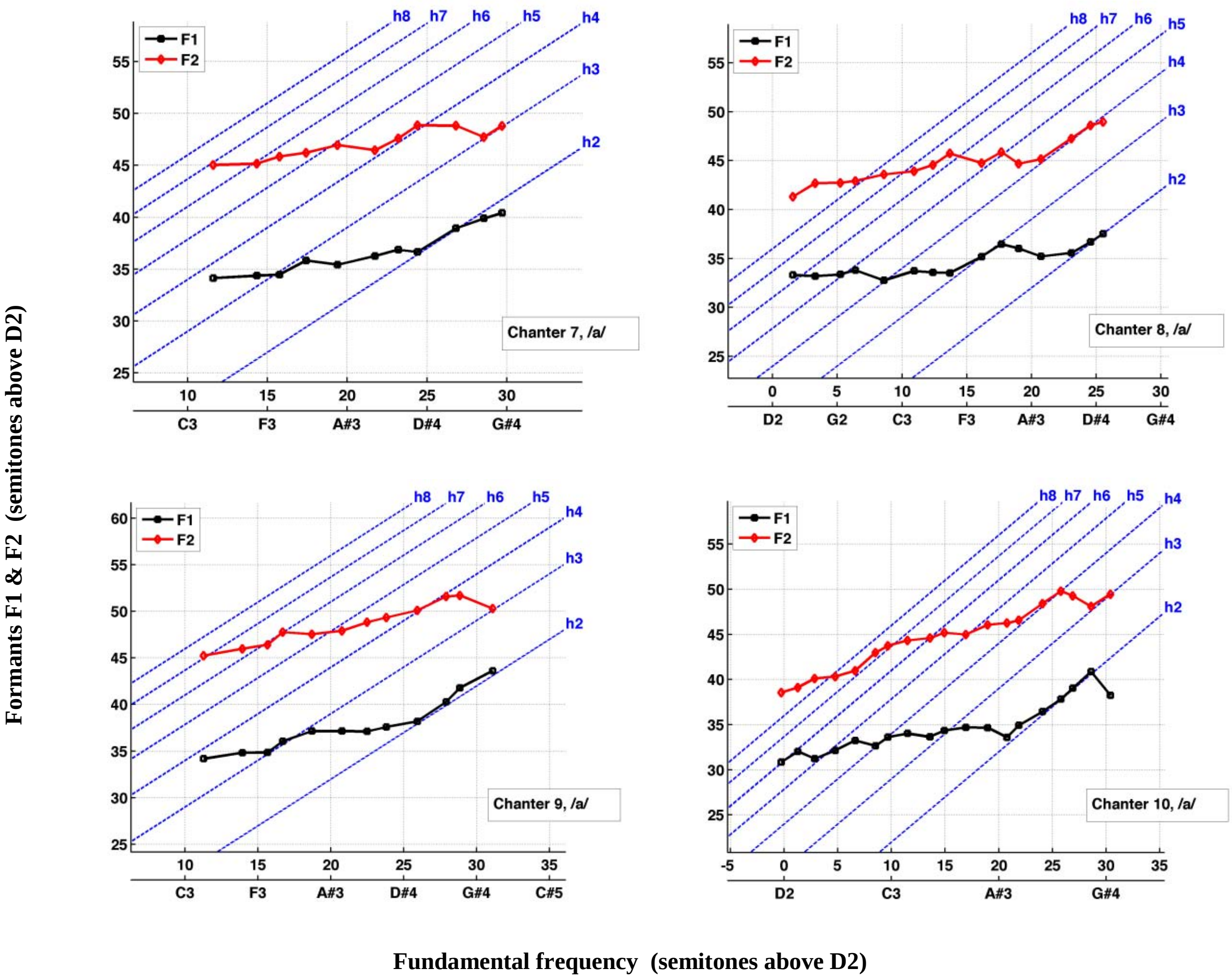

Figure 5 (continued)

Chanter 8 seems to start tuning its F1 formant to h3 at about E3 and then tunes it to h2 at G#3. Its F2 formant remains relatively constant until A3 where it tunes to h4.

Chanter 9 mainly tunes F1 to h2 at E4, but also shows signs of tuning F1 to h3 at the region between F#3 and A3. Its F2 formant tunes to h4 above E4 and then jumps to h3 at the last note of its range.

Chanter 10 is the clearest evidence of formant tuning among the sample of 10 chanters. It keeps its F1 tuned to h2 throughout the range of B3-G4 that is 6 consecutive notes. In the same pitch region F2 is tuned to h4 for the first 5 notes then finally tunes to h3.

5. DISCUSSION

Previous research has shown that singing voice classification could be based on the formant frequency values and vocal range [30, 31]. This way a chanting voice can be classified in one of the classical music voice types. Moreover, although BCM performers have not been exposed, traditionally, to classical vocal training techniques, formant tuning can be considered as an intuitive act [4], and therefore be performed by non-classical trained voices as well.

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 [7, 10]. It has been found that formant tuning is mostly observed in and above this range [10]. 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, 10], as well as the maximum distance criterion of 50 Hz between the formant and its nearest harmonic, used in similar works [10], 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 [10]. 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 the vowel /a/, is displayed in **Figure 5**. As before, the data presented in the graph are for the ascending diatonic scale, for all ten chanters (Ch1-Ch10). In order to draw safer conclusions, only notes above C4 (22 semitones above D2) have been accounted for. Indeed, the particular pitch range is near the passaggio region, where formant tuning is more likely to occur.

By examining the F1 graph of **Figure 6** we notice that at least 6 of the total of 10 chanters, analyzed in the current investigation, tend to tune their F1 to the h2 harmonic, at the pitch range above C4.

Formant tuning was most apparent in cases where the F1 and F2 remained relatively constant throughout the scale, before reaching a breaking point near the beginning of the passaggio region, like in the cases of chanters 1, 2, 8, 9 and 10. Another observation is that F2 formant of the ultimate scale’s note was tuned to either h3 or h4 in all chanters. In fact the F2 graph of **Figure 6** shows that, above C4, all chanters seem to have their F2 tuned to one of the h3, h4 partials.

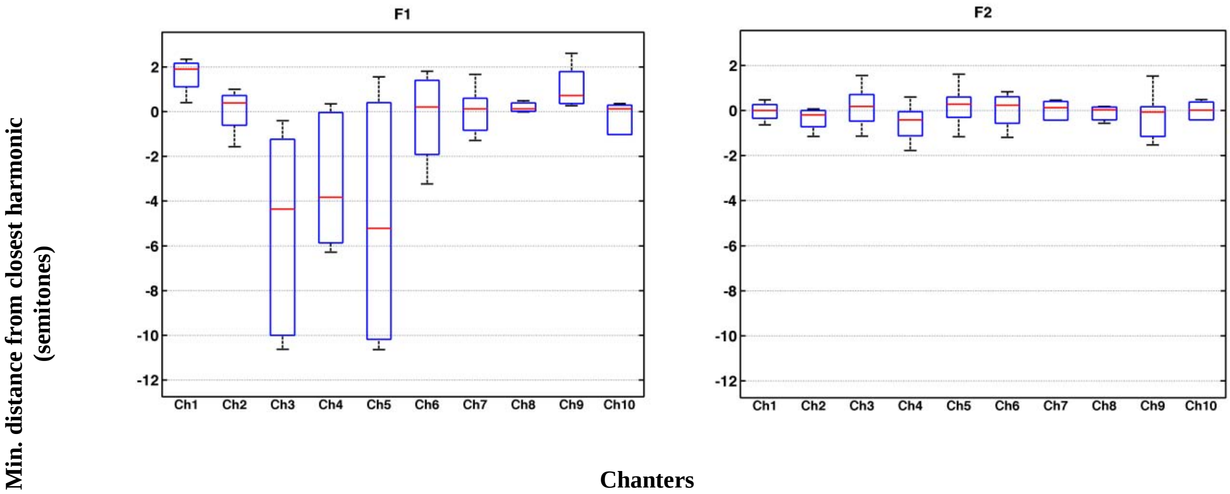

Figure 6: Distance in semitones between the first formant F1 and its closest harmonic and between the second formant F2 and its closest harmonic, for the vowel /a/. Results of all ten chanters (*Ch1-Ch10*) are presented for the ascending diatonic scale, for pitches above C4.

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. This last assumption agrees with the answer given by Sundberg et al. [10] and Henrich et al. [32] to a similar question, for the passaggio transition. Our opinion is also supported by the fact that most professional Byzantine chanters usually have not been given a systematic phonetic training as part of their vocal pedagogy that could lead in standardizing the Byzantine chant vocal technique.

6. CONCLUSIONS

It is evident from both **Figure 2** and **Figure 6** that formant tuning is in use by the modern Byzantine chant performers. Not only at high f0 but also at the lower and middle scale regions, formant tuning has been observed in most of the total 10 chanters analyzed in the current investigation.

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

A closer examination of the vocal range above C4 reveals that all chanters, in our sample, had their F2 resonance tuned to the partial h3 or h4.

In contrast to other types of singing voice, our results cannot be directly compared to other findings for the BCM in the field of formant tuning, since, to our knowledge, there aren’t any similar publications on this subject. However, when compared to results of professional male opera singers [10] from recent investigations, one might conclude that there is a stronger formant tuning tendency in BCM.

Acknowledgments

This research 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.

REFERENCES

[1] Y. Hu, D. Wu, and A. Nucci, “Pitch-based gender identification with two-stage classification,” *Security and Communication Networks*, vol. 5, no. 2, pp. 211–225, 2012.
[2] G. Fant, *Acoustic Theory of Speech Production*. Walter de Gruyter, 1970.
[3] G. Carlsson and J. Sundberg, “Formant frequency tuning in singing,” *J. of Voice*, vol. 6, no. 3, pp. 256–260, 1992.
[4] J.A. Stark, *Bel canto: a history of vocal pedagogy*. University of Toronto Press, 2003.
[5] J. Sundberg, “The acoustics of the singing voice,” *Sci. Am.*, vol. 236, no. 3, pp. 82–91, 1977.
[6] J. Sundberg, *The Science of the Singing Voice*. Northern Illinois Univ, 1987.
[7] D.G. Miller and H.K. Schutte, “Formant tuning in a professional baritone,” *J. of Voice*, vol. 4, no. 3, pp. 231–237, 1990.
[8] G. Carlsson and J. Sundberg, “Formant frequency tuning in singing,” *J. of Voice*, vol. 6, no. 3, pp. 256–260, 1992.
[9] 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.
[10] J. Sundberg, F.M.B. Lã, and B.P. Gill, “Formant Tuning Strategies in Professional Male Opera Singers,” *J. of Voice*, Dec. 2012.
[11] 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.
[12] 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.
[13] D.G. Panagiotopoulos, *Theoria kai Praxis tis Byzantinis Ecclesiastikis Mousikis*, 6th ed. Athens: Sotir, 1997.
[14] 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.
[15] M. Chrysanthos and P.G. Pelopides, *Theoretikon mega tes mousikes*. En Tergeste : ek tes typographias Michael Vais (Michele Weis), 1832.
[16] 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.

[17] 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.
[18] 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.
[19] P. Boersma, "Praat, a system for doing phonetics by computer," *Glot International*, vol. 5, no. 9/10, pp. 341–345, 2002.
[20] 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.
[21] N. Andersen, "On the calculation of filter coefficients for maximum entropy spectral analysis," *Geophysics*, vol. 39, no. 1, pp. 69–72, Feb. 1974.
[22] D. G. Childers, *Modern spectrum analysis*. IEEE Press : sole worldwide distributor (exclusive of IEEE), Wiley, pp. 252-255, 1978.
[23] 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.
[24] 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.
[25] T. Wempe and P. Boersma, "The interactive design of an F0-related spectral analyser," in *Proc. 15th ICPhS*, 2003, pp. 343–346.
[26] 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.
[27] 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.
[28] 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.
[29] 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.
[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.

Georgios Kouroupetroglou *holds a B.Sc. in Physics and a Ph.D. in Communications and Signal Processing. Member of the Academic staff of the Division of Communication and Signal Processing, Department of Informatics and Telecommunications, National and Kapodistrian University of Athens and leader of the Speech and Accessibility Laboratory. His research interests focus on the area of Voice User Interfaces and Computer Accessibility, as a part of the major domain of Human-Computer Interaction, with emphasis in: Spoken Dialogue Human Computer Interaction, VoiceWeb, Voice Agents, Accessible Computing, Voice Processing; Analysis and Synthesis of Speech and Singing, Music Computing; Singing Voice Analysis, Virtual Musical Instruments, and Byzantine Chanting Analysis and Synthesis. Professor Kouroupetroglou has actively participated in a number of European Union funded and National research projects. He has been a reviewer / evaluator and member of working groups/technical panels of various European Union's projects/programs.*

George Chrysochoidis *holds an MSc in Digital Signal Processing from the department of Informatics & Telecommunications at the National & Kapodistrian University of Athens and is currently pursuing a PhD* *program at the same department. As a freelancer, he has worked in the field of custom software development as well as providing IT consultancy services for companies. In 2005 he joined the signal processing team in the department of Informatics & Telecommunications at the National & Kapodistrian University of Athens and since then has worked on several projects. His research interests include speech recognition, singing voice analysis, modeling and synthesis, Byzantine Chant analysis and synthesis, natural language processing and machine learning.*