

Σύγκριση των φωνηέντων της Βυζαντινής Μουσικής με τα αντίστοιχα της ομιλίας και του ανδρικού Οπερατικού τραγουδιού

Δελβινιώτης Δημήτριος (ddelvis@otenet.gr)
Θεοδωρίδης Σέργιος (Stheodor@di.uoa.gr)
Αμαργιαννάκης Γεώργιος

Τμήμα Πληροφορικής
Πανεπιστήμιο Αθηνών
Πανεπιστημιούπολις, Κτίρια ΤΥΠΑ
15784 Αθήνα
<http://www.di.uoa.gr>

Περίληψη

Επειδή οι μελέτες στην βυζαντινή μουσική, όσον αφορά στην ακουστική ανάλυση των ψαλτικών ήχων, είναι ελάχιστες έως ανύπαρκτες, η παρούσα μελέτη καταγίνεται με μια πρώτη σύγκριση μεταξύ των φωνοσυντονισμών (formants) των φωνηέντων της ψαλτικής τέχνης με τα αντίστοιχα της ομιλίας στην Ελληνική γλώσσα καθώς και με τα παρόμοια του ανδρικού οπερατικού τραγουδιού. Η μόνη παράμετρος των formants που εξετάζεται είναι η συχνότητα.

1. Εισαγωγή

Ο όρος «Βυζαντινή Μουσική», ευρύτατα γνωστός, συχνά συνοδεύεται από τον επιθετικό προσδιορισμό «Εκκλησιαστική» για να δηλώσει την διαφορά του από την ούτω λεγόμενη «Κοσμική ή Εξωτερική» μουσική. Οι ιεροψάλτες ή απλώς ψάλτες και οι ιερείς είναι οι φυσικοί θεραπευταί της Ψαλτικής Τέχνης. Η βυζαντινή μουσική είναι μια καθαρώς φωνητική μονωδική τέχνη και οργανικά ασυνόδευτη¹ ενώ συνοδεύεται από το ισοκράτημα, το οποίο συνίσταται στην ταυτόχρονη συνήχηση με την κύρια μελωδία της βασικής νότας-φθόγγου (τονικής) του ήχου εντός του οποίου κινείται η μελωδία². Συχνά το ισοκράτημα είναι μια οκτάβα κάτω από την τονική του ψαλλομένου ήχου και ασθενές ως προς την ένταση. Με τον όρο «ήχος» εννοούμε ένα σύνολο μουσικών χαρακτηριστικών βάσει των οποίων διακρίνομε την μία μελωδία από την άλλη: «Ήχος είναι ιδέα μελωδίας, συνισταμένη εις τήν ἑξίν του γινώσκειν, τίνας μεν των φθόγγων αφετέον, τίνας δε παραληπτέον· και από τίνος τε αρτέον, και εις ὃν καταληκτέον»³. Τα ψαλλόμενα τροπάρια, δηλαδή τα μελωδήματα της Εκκλησίας⁴, ἄδονται σε τονική βάση βολική για τον ψάλτη ώστε να αισθάνεται άνετα κατά την ψαλμωδία. Επειδή συνήθως η μουσική έκταση της μελωδίας τους είναι περίπου μία οκτάβα, οι νότες που ακούγονται συχνότερα μεταξύ διαφορετικών ψαλτών είναι μεταξύ του C3(132Hz) και του C4(264Hz)⁵. Αν και υπάρχουν διάφορες απόψεις περί του ύφους φαίνεται ότι μάλλον έχουν διαμορφωθεί ποικίλες σχολές ερμηνείας των μελωδιών⁶.

Παρ' όλη την έλλειψη ακουστικών μελετών στην βυζαντινή μουσική, το τραγούδι στην όπερα όπως επίσης και σε άλλα μουσικά στυλ έχει αρκούντως μελετηθεί. Σημαντικό μέρος των μελετών αφορούν τα formants (φωνοσυντονισμοί) τα οποία είναι οι συντονισμοί της φωνητικής οδού (vocal tract) στην πηγάα ηχητική διέγερση (voice source) που παράγεται στην γλωττίδα, και χαρακτηρίζονται από τρία στοιχεία: 1) την κεντρική συχνότητα, 2) το εύρος ζώνης, και 3) την στάθμη ισχύος των⁷. Έχει βρεθεί ότι τα δύο πρώτα formants F1 και F2 είναι υπεύθυνα για τήν φωνηεντική ποιότητα⁸, δηλαδή την

¹ ΤΑ ΠΑΘΗ ΤΑ ΣΕΠΤΑ, Σειρά δίσκων. Έκδοση Ελληνικού Πολιτιστικού Κέντρου Λονδίνου - Επιμέλεια: Γρηγόριος Στάθης, 1982, σελ.33.

² ΜΕΘΟΔΟΣ ΤΗΣ ΕΛΛΗΝΙΚΗΣ ΜΟΥΣΙΚΗΣ ΤΟΜ. Β', Σίμωνος Καρά, Αθήνα, 1982, σελ. 200

³ ΘΕΩΡΗΤΙΚΟΝ ΜΕΓΑ ΤΗΣ ΜΟΥΣΙΚΗΣ, Χρυσάνθου Αρχιεπισκόπου Δυρραχίου εκδοθέν υπό Π.Γ. Πελοπίδου εν Τεργέστη 1832, σελ. 125

⁴ ΠΑΡΗΧΗΜΑΤΑ, Γρηγ. Στάθης, Αθήνα 1978, σελ. 106

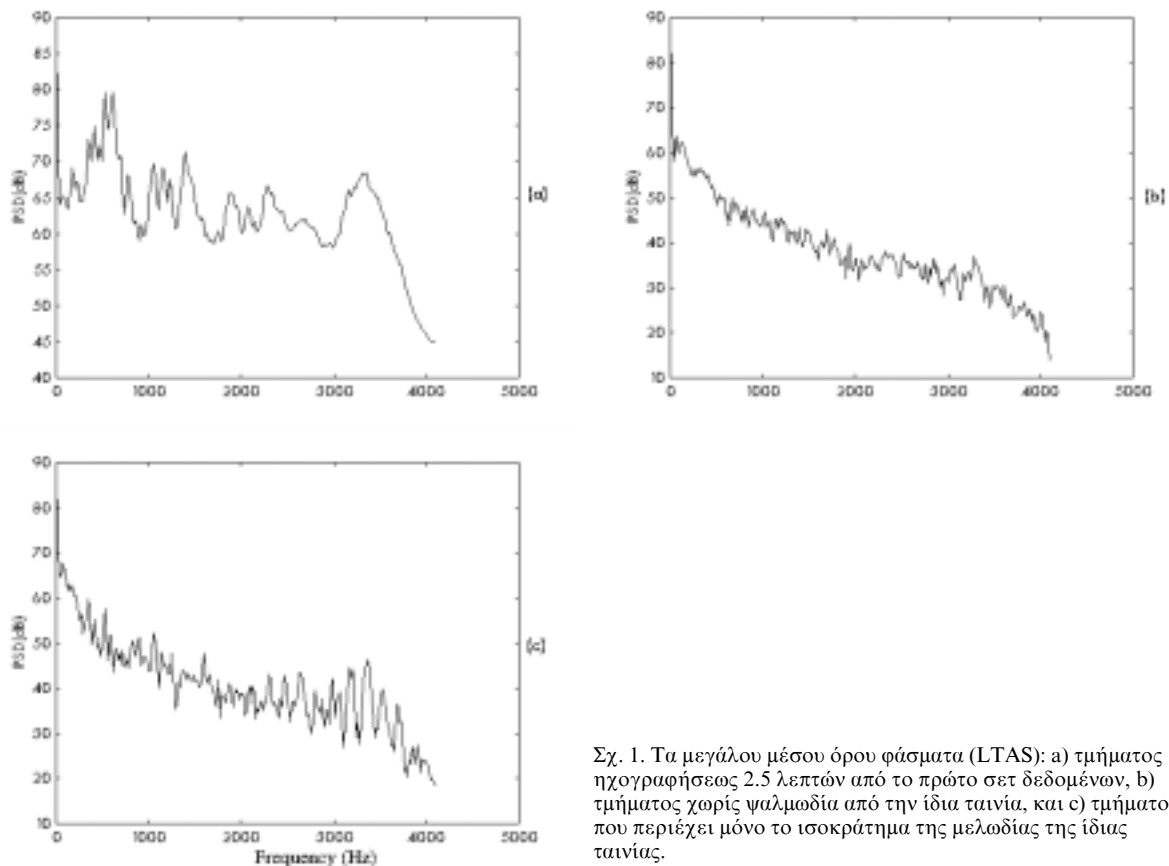
⁵ ΕΙΣΑΓΩΓΗ ΣΤΗ ΦΥΣΙΚΗ ΚΑΙ ΜΟΥΣΙΚΗ ΑΚΟΥΣΤΙΚΗ, Χαράλαμπος Χ. Σπυρίδης, Αθήνα, 1999, σελ. 88-96.

⁶ ΒΥΖΑΝΤΙΝΗ ΕΚΚΛΗΣΙΑΣΤΙΚΗ ΜΟΥΣΙΚΗ ΚΑΙ ΨΑΛΜΩΔΙΑ, ΤΟΜ. Α', Φίλιππος Α. Οικονόμου, Αίγιο, 1992, σελ.245.

⁷ DISCRETE-TIME PROCESSING OF SPEECH SIGNALS, J. R. Deller, Jr - J. G. Proakis - J. H. L. Hansen, USA 1993, σελ. 101 - 140.

⁸ THE SCIENCE OF THE SINGING VOICE, Johan Sundberg, Northern Illinois 1987, σελ. 20 - 24.

διάκριση μεταξύ των φωνηέντων, ενώ τά ανώτερα και κυρίως τα F3, F4, και F5 παίζουν σημαντικό ρόλο στο ανδρικό οπερατικό τραγούδι⁹. Η συχνότητα του F1 είναι μικρότερη στο τραγούδι



Σχ. 1. Τα μεγάλου μέσου όρου φάσματα (LTAS): a) τμήματος ηχογραφήσεως 2.5 λεπτών από το πρώτο σετ δεδομένων, b) τμήματος χωρίς ψαλμωδία από την ίδια ταινία, και c) τμήματος που περιέχει μόνο το ισοκράτημα της μελωδίας της ίδιας ταινίας.

συγκρινόμενη με την αντίστοιχη της ομιλίας για τα περισσότερα φωνήεντα. Οι συχνότητες των F2 και F3 είναι επίσης χαμηλότερες για τα εμπρόσθια (front) φωνήεντα¹⁰ αλλά η συχνότητα του F3 μεγαλώνει για τα οπίσθια φωνήεντα¹¹. Για όλα τα φωνήεντα, τα formants F3, F4 και F5 συγκεντρώνονται έτσι ώστε να δημιουργείται το formant του τραγουδιστή (singer's formant)¹² γεγονός που συνοδεύεται με την κάθοδο του λάρυγγα και είναι το αποτέλεσμα της διευρύνσεως του φάρυγγα έτσι ώστε να γίνει 6 φορές το άνοιγμα του λάρυγγα. Επίσης όταν ο λάρυγγας κατεβαίνει χαμηλώνουν οι συχνότητες όλων των φωνοσυντονισμών όπως και όταν προτείνονται τα χείλη. Το σχήμα του σώματος της γλώσσας ρυθμίζει την συχνότητα του F2 η οποία κατεβαίνει όταν ο όγκος της γλώσσας πηγαίνει προς τον φάρυγγα (οπίσθια φωνήεντα) και ανεβαίνει όταν πλησιάζει το εμπρόσθιο τμήμα της στοματικής κοιλότητας (εμπρόσθια φωνήεντα). Όσο μεγαλύτερη είναι η φαρυγγική στένωση εξ αιτίας της γλώσσας τόσο χαμηλότερη είναι η συχνότητα του F2¹³.

Το ηχοχρώμα της φωνής ενός οπερατικού τραγουδιστή, αν βασιστούμε στην αντίληψή μας, ακούγεται σαφώς διαφορετικό από εκείνο ενός τυπικού ψάλτη. Καθώς το φωνητικό ηχοχρώμα καθορίζεται από τις συχνότητες των φωνοσυντονισμών της φωνητικής οδού, είναι καθοριστικής σημασίας η ανάλυση και μελέτη του. Έτσι θα κατανοηθεί τόσο η άρθρωση (articulation) του φωνητικού οργάνου για την παραγωγή των ψαλτικών ηχοχρωμάτων όσο και οι τιμές των παραμέτρων των φωνητικών ακουστικών μεγεθών τα οποία περιγράφουν αυτά τα ηχοχρώματα.

Στην παρούσα ανακοίνωση καταγράφουμε τις συχνότητες των φωνοσυντονισμών για ένα μεγάλο αριθμό ψαλτικών ηχοακουσμάτων και τις συγκρίνουμε με εκείνες της ομιλίας και των αντιστοίχων οπερατικών. Στην παράγραφο 2 παρουσιάζουμε τα σετ δεδομένων που αναλύσαμε στην έρευνά μας. Στην παράγραφο 3 περιγράφουμε τα εργαλεία επεξεργασίας σήματος (signal processing) που χρησιμοποιήσαμε και στην τελευταία παράγραφο παρουσιάζουμε τα αποτελέσματα.

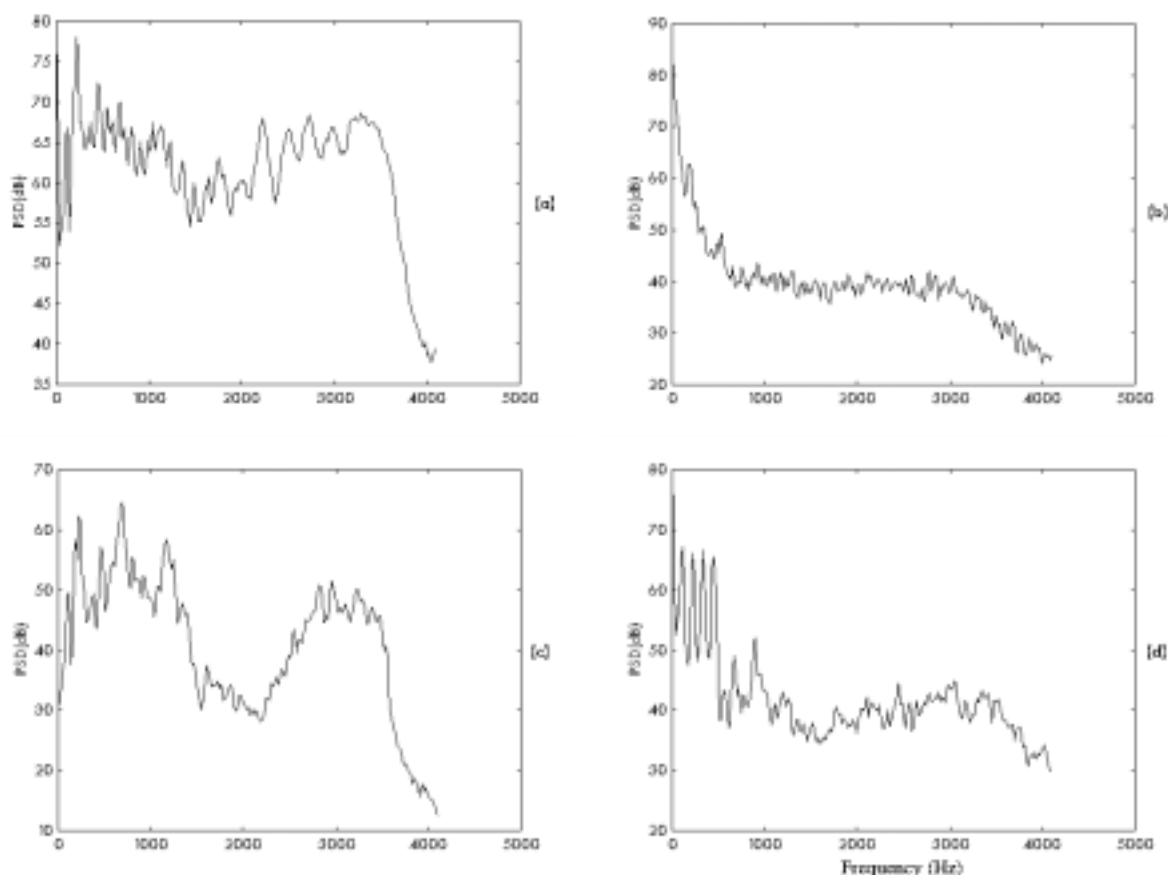
⁹ βλ. ανωτέρω The science of the singing voice, σελ. 115 - 123.

¹⁰ Τα φωνήεντα ομαδοποιούνται σε τρεις κατηγορίες ανάλογα με την θέση του όγκου της γλώσσας μέσα στην στοματική κοιλότητα: εμπρόσθια - κεντρικά - οπίσθια.

¹¹ Formant structure and articulation of spoken and sung vowels. Folia Phoniatrica 22: 28 - 48

¹² Ιδέ ανωτέρω άρθρο υποσημείωσης 11.

¹³ Ιδέ υποσημ. 7, σελ. 124.



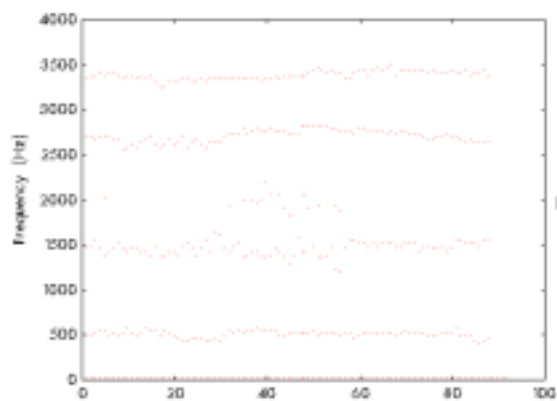
Σχ. 2. Τα LTAS από: a) τμήμα ηχογράφησης του δευτέρου σετ δεδομένων, b) τμήματος χωρίς ψαλμωδία, c) τμήματος μόνο με το φωνήεν /a/ και, d) μόνο το ισοκράτημα του ίδιου τμήματος ηχογραφήσεως.

2. Υλικό

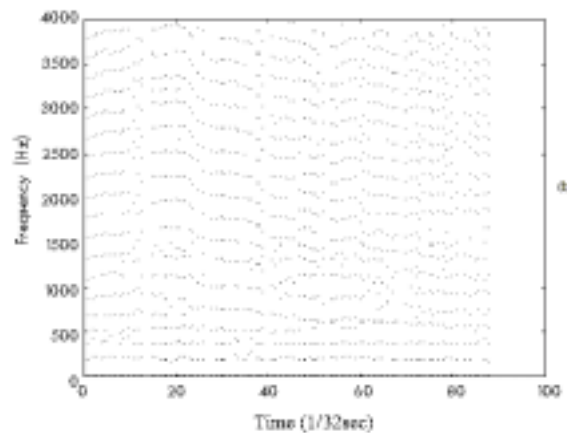
Στην έρευνά μας χρησιμοποιήσαμε τρία σετ δεδομένων. Το πρώτο αποτελείται από ένα μεγάλο αριθμό ψαλμωδιών από ποικίλους ψάλτες καταγεγραμμένων μέσα σε μεγάλη χρονική περίοδο που αρχίζει από τις αρχές του αιώνα μας και φθάνει μέχρι σήμερα. Αυτό το υλικό ανήκει σε προσωπικό αρχείο των συγγραφέων. Ο μέσος όρος του λόγου σήμα προς θόρυβο (signal to noise ratio - SNR) γι' αυτό το σετ ήταν περίπου 10dB για την περιοχή του φάσματος 0 - 2kHz και 25dB για την συχνοτική περιοχή 2 - 4kHz με τυπικές αποκλίσεις τιμών περίπου 3dB και 7dB αντίστοιχα. Τά σχήματα 1(a) - (c) δείχνουν τρία τυπικά φάσματα απ' αυτό το σετ δεδομένων τα οποία αντιστοιχούν σε ένα κομμάτι ψαλμωδίας, στο θόρυβο του κομματιού και στο ισοκράτημά του. Παρόλο που το υλικό αυτό είναι αντιπροσωπευτικό ως προς τα διάφορα ψαλτικά στυλ που χρησιμοποιήθηκαν μέχρι τώρα, οι συνθήκες καταγραφής του δεν είναι σαφώς ελεγχόμενες και για αυτόν τον λόγο χρειάζεται περαιτέρω έρευνα και επαλήθευση. □Ετσι δημιουργήσαμε άλλα δύο σετ δεδομένων με καλά καθορισμένες συνθήκες καταγραφής.

Το δεύτερο σετ δεδομένων περιέχει ψαλμωδήματα από έμπειρους ψάλτες τα οποία κατεγράφησαν σε θάλαμο ηχογραφήσεων σε μικρή απόσταση από το μικρόφωνο (περίπου 0,5m εμπρός από το στόμα) μέσω επαγγελματικού μαγνητοφώνου. Οι λόγοι SNR στην χαμηλή και υψηλή φασματική περιοχή ήσαν 25dB και 30dB αντίστοιχα. Το ισοκράτημα συνόδευε όλες τις ηχογραφήσεις και είχε στάθμη ήχου περίπου 2dB πάνω από τον θόρυβο, για την χαμηλή περιοχή, και 3dB για την υψηλή περιοχή (βλ. σχ. 2(a) - (d)). Στο σχήμα 2(d) υπάρχει μια πολύ χαμηλή συχνοτική περιοχή (0 - 500Hz) όπου το ίσον έχει σχετικά υψηλή στάθμη (περίπου 5 - 10dB) αλλά αυτό δεν εμποδίζει την εκτίμηση του F1 formant για τον λόγο ότι η συχνότητα του ίσου δεν μεταβάλλεται σύμφωνα με εκείνη του F1 ενώ παράλληλα η συχνότητα του F1 δεν είναι ίση με καμιά αρμονική του ίσου (σχ. 3). Βάσει αυτού του σετ δεδομένων υπολογίστηκαν οι συχνότητες των formants και συγκρίθηκαν τα αποτελέσματα με εκείνα του πρώτου σετ.

Οι ηχογραφήσεις από τρεις διαφορετικούς ψάλτες, όπου για τον καθένα έγιναν καταγραφές κάτω από τρεις διαφορετικές συνθήκες, συνιστούν το τρίτο σετ δεδομένων. Πιο συγκεκριμένα μια ομάδα καταγραφών έγινε σε ανηχώικό δωμάτιο με μικροφωνική απόσταση περίπου 0.5m, μια άλλη ομάδα έγινε μέσα σε εκκλησία με



Σχ. 3. a) Τροχιές formants (μέθοδος Markel) για το φωνήεν /e/ ψαλτικού ήχου.



Σχ. 3. b) Τροχιές των αρμονικών για το ίδιο φωνήεν και ήχο. Οι πρώτες 2 αρμονικές του ήσου σχηματίζουν ευθείες γραμμές.

αντηχήσεις (μιμούμενοι κάποιες απ' τις καταγραφές του πρώτου σετ) στην ίδια μικροφωνική απόσταση, και μια τρίτη ομάδα καταγραφών έγινε μέσα στην ίδια εκκλησία σε μεγαλύτερη μικροφωνική απόσταση (περίπου 1m). Ο σκοπός αυτού του σετ ήταν να προσδιορισθεί η επίδραση των διαφορετικών συνθηκών στην εκτίμηση των φορμαντικών συχνοτήτων του πρώτου σετ. Βρέθηκε ότι η μέση διακύμανση για μια τιμή συχνότητας (variation) για τον ίδιο ψάλτη αλλά σε διαφορετικές συνθήκες είναι πολύ μικρότερη εν σχέσει προς εκείνη που βρέθηκε κάτω απ' την ίδια συνθήκη αλλά για διαφορετικούς ψάλτες. □ Έτσι οι μέσες τιμές φορμαντικών συχνοτήτων για ένα σύνολο ψαλτών επηρεάζονται κυρίως από τις μεταξύ τους διαφορές ενώ οι αλλοιώσεις στις ανωτέρω τιμές λόγω διαφορετικών συνθηκών καταγραφής είναι μη υπολογίσιμες για τον σκοπό αυτής της μελέτης.

Τέλος, ως προς το είδος των μελωδημάτων που επιλέξαμε για την ανάλυση, το καταλληλότερο ήταν το αργό ή το αργοσύntonτο έτσι ώστε να περιέχουν μακράς διάρκειας μουσικούς φθόγγους - νότες σε κάθε φωνήεν και τονικό ύψος (pitch). Ο λόγος ήταν για να διατηρείται σε κάθε φθόγγο μια σταθερή άρθρωση έτσι ώστε να εξαχθούν συμπεράσματα για ένα πιθανό προφίλ του φωνητικού οργάνου που υπάρχει στην ψαλτική φωνή. Τό τονικό ύψος των μελωδημάτων επελέγη μεταξύ των C3 και C4 ώστε να είναι αυτά που συνήθως ακούγονται μέσα στους ναούς και να έχουν παραχθεί χωρίς βεβιασμένη φωνή.

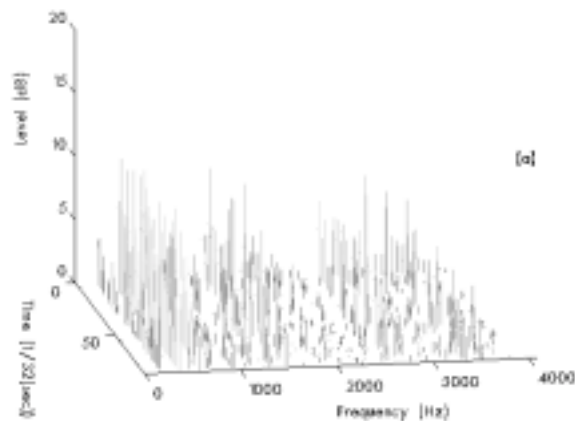
3. Μέθοδος ανάλυσης

Από όλα τα σετ δεδομένων εξήχθησαν ηχητικά τμήματα, για κάθε ψάλτη, χαρακτηριζόμενα απ' τα εξής: 1) Την ύπαρξη φθόγγων της μεσαίας φωνητικής περιοχής, δηλαδή μεταξύ των C3 και C4, 2) νότες σχετικά μακράς διάρκειας (μέχρι 2sec), 3) την απουσία ουσιαστικών αλλαγών στην ηχητική ένταση, 4) την παρουσία όλων των Ελληνικών φωνηέντων /α, ε, ι, ο, υ/ (IPA symbols), και 5) την όσο το δυνατόν παρουσία περισσότερων τονικών υψών για κάθε φωνήεν.

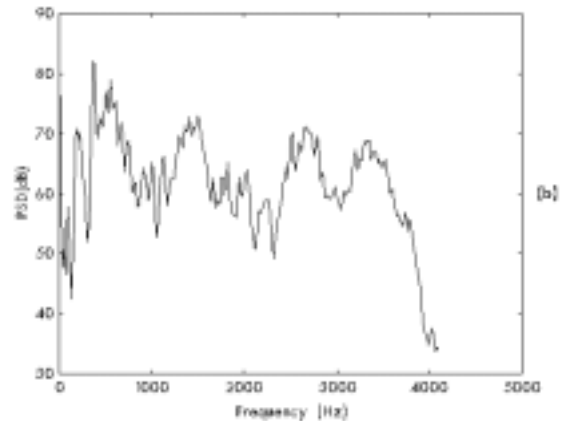
Οι ηχογραφήσεις ψηφιοποιήθηκαν με συχνότητα δειγματοληψίας 8192Hz στα 16bits/δείγμα. Μετά υπολογιζόταν η φασματική πυκνότητα ισχύος PSD (Power Spectral Density) για κάθε ηχητικό τμήμα που περιέχει ένα φωνήεν σε όσο το δυνατόν περισσότερες νότες. Ο λόγος που δεν χρησιμοποιήσαμε μεγαλύτερη συχνότητα δειγματοληψίας ήταν διττός. 1) Η ενέργεια του φάσματος είναι πολύ μικρή πέραν των 4kHz και 2) η ταχύτητα επεξεργασίας από τις ρουτίνες ανάλυσης έπρεπε να είναι αρκετή για την επεξεργασία τόσου μεγάλου πλήθους δεδομένων (186*5=930 ηχητικά τμήματα). Η μέθοδος εύρεσης της PSD ήταν η μέθοδος Welch η οποία βρίσκει τον μέσο όρο (averaging) φασμάτων (περιοδογραμμάτων - periodograms) διαδοχικών ή εν επικαλύπτει μπλόκ-τμημάτων στα οποία κόβεται το υπό ανάλυση τμήμα¹⁴. Τα μπλόκ δημιουργούνται με κόψιμο του αρχικού τμήματος μέσω παραθύρου Hanning με μήκος 512 δείγματα και επικάλυψη μεταξύ τους 50%. Το τελικό φάσμα μεγάλου χρονικού μέσου όρου (LTAS - Long - Time Average Spectrum) (βλ. σχ. 4(b)) δείχνει χονδρικά την παρουσία τεσσάρων κορυφών που αντιστοιχούν στα τέσσερα πρώτα formants.

¹⁴ DIGITAL SPECTRAL ANALYSIS WITH APPLICATIONS, S. Lawrence Marple, Jr. Prentice-Hall Signal Processing Series, New Jersey 1987, σελ. 154

Η εκτίμηση των φορμαντικών συχνοτήτων έγινε με την μέθοδο Markel για κάθε ηχητικό τμήμα, όπου κατασκευάζεται ένα κατάλληλο φίλτρο ώστε να αναιρούνται οι επιδράσεις της φωνητικής οδού δηλαδή οι



Σχ. 4 α) Το 3-διάστατο σπεκτρογράμμο για τον ήχο του σχ. 3 βάσει της ανάλυσης AR όπου φαίνεται και η ισχύς κάθε αρμονικής



Σχ. 4 β) Το αντίστοιχο LTAS του ίδιου ψαλτικού ήχου με τα σχήματα 4α, 3α και 3β.

φωνοσυντονισμοί (inverse filtering). Τό ψηφιακό φίλτρο σχεδιάζεται έτσι ώστε στην έξοδο του να ελαχιστοποιείται η ολική ενέργεια του σήματος που μπαίνει στην είσοδό του. Ο αριθμός M των παραμέτρων του φίλτρου είναι η τάξη του φίλτρου και στην ανάλυσή μας επελέγη $M=13$. Το μήκος των μπλοκ προς ανάλυση ήταν 256 δείγματα. Τα ελάχιστα της συναρτήσεως αποκρίσεως συχνότητας του φίλτρου αντιστοιχούν

TABLE I

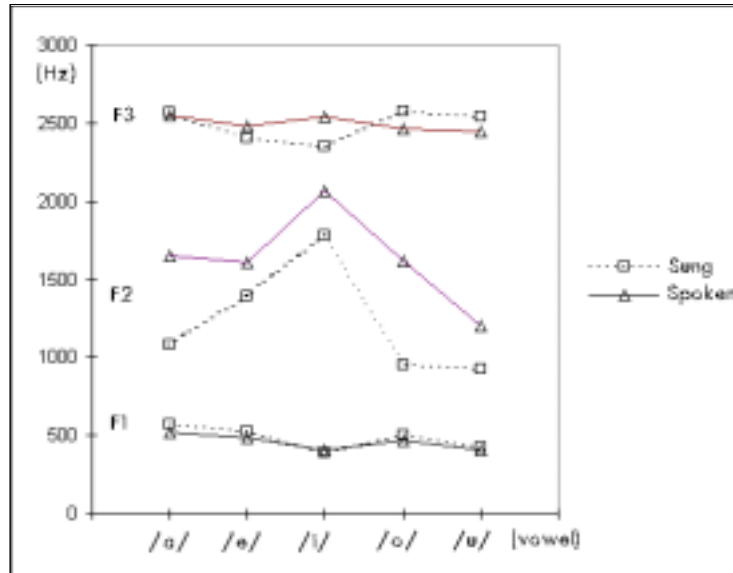
Οι μέσες φορμαντικές συχνότητες και οι ισχύες τους με τις αντίστοιχες τυπικές τους αποκλίσεις για το πρώτο και δεύτερο σετ ηχητικών τμημάτων, για κάθε φωνήεν. Οι μέσες τιμές συχνοτήτων είναι μεταξύ των δύο σετ συγκρίσιμες ενώ οι τυπικές αποκλίσεις του πρώτου σετ είναι γενικά μεγαλύτερες εκείνων του πρώτου σετ. Οι μέσες τιμές των ισχύων τους είναι επίσης συγκρίσιμες μεταξύ των δύο σετ.

Vowel \ Formant	F1 (Hz)	sdF1	F2 (Hz)	SdF2	level_F1	sd_levelF1	level_F2	sd_levelF2	
/a/	574	79	1083	105	0	2	-4	5	First
/e/	530	76	1395	146	-1	2	-7	8	Set
/i/	387	78	1785	173	-1	3	-9	10	
/o/	499	75	949	107	0	1	-5	7	
/u/	421	80	925	157	0	2	-8	8	
/a/	500	77	1014	76	-1	1	-4	4	Second
/e/	466	37	1356	136	-1	2	-6	6	Set
/i/	349	35	1674	144	-1	2	-10	7	
/o/	420	49	869	80	-1	1	-5	5	
/u/	362	47	817	127	0	0	-7	5	
Vowel \ Formant	F3 (Hz)	sdF3	F4 (Hz)	SdF4	level_F3	sd_levelF3	level_F4	sd_levelF4	
/a/	2580	248	3228	200	-18	10	-18	10	First
/e/	2406	172	3218	204	-16	10	-18	10	Set
/i/	2364	189	3221	201	-13	11	-16	12	
/o/	2590	266	3210	198	-21	11	-20	12	
/u/	2556	272	3200	190	-23	13	-22	12	
/a/	2529	242	3064	141	-16	8	-12	7	Second
/e/	2308	176	2974	159	-12	7	-9	6	Set
/i/	2212	180	2985	159	-14	8	-11	8	
/o/	2501	241	3039	135	-22	11	-14	7	
/u/	2475	244	3014	155	-19	9	-15	8	

στους φωνοσυντονισμούς της φωνητικής οδού¹⁵. Συνοπτικά η μέθοδος έχει ως εξής: Για την σειρά των μπλοκ, στα οποία τεμαχίζεται το προς ανάλυση ηχητικό τμήμα, υπολογίζονται οι φωνοσυντονισμοί για κάθε μπλοκ και σχεδιάζονται με την χρονική τους διαδοχή σε ένα σχήμα συχνότητας - χρόνου (σχ. 3(a)). Η εκτίμηση των formants για το συνολικό ηχητικό τμήμα γίνεται με σύγκριση των τροχιών που σχηματίζουν με την τροχιά της θεμελιώδους συχνότητας οι οποίες πρέπει να διαφέρουν σημαντικά μεταξύ των. Αν η τροχιά κάποιου formant μοιάζει με την τροχιά της θεμελιώδους τότε η μέθοδος κάνει

¹⁵ Digital Inverse Filtering - A new tool for formant trajectory estimation, John D. Markel, IEEE TRANSACTIONS ON AUDIO AND ELECTROACOUSTICS, 20, 129-137, 1972

λανθασμένη εκτίμηση για το συγκεκριμένο formant έχοντας εγκλωβιστεί σε μια ισχυρή αρμονική και έτσι αυτό το formant πρέπει να αγνοηθεί. Επίσης μια κορυφή μεγάλου εύρους στο αντίστοιχο LTAS φάσμα στην συχνότητα που ανήκει μια συγκεκριμένη τροχιά επιβεβαιώνει την παρουσία ενός formant στο σημείο αυτό (σχ. 3(a), 4(b)). Από το σύνολο των τιμών συχνότητας μιας φορμαντικής τροχιάς επιλέγεται εκείνη που εμφανίζεται συχνότερα, έστω και με μικρή διακύμανση γύρω από μια μέση τιμή επειδή αυτή η τιμή γίνεται αντιληπτή από το ανθρώπινο αυτί¹⁶.



Σχ. 5 Οι τρεις πρώτες μέσες φορμαντικές συχνότητες για την ομιλία και ψαλμωδία ανά φωνήεν. Είναι φανερή η ελαττωμένη τιμή της συχνότητας του F2 formant καθώς και η ελαφρά αυξημένη τιμή της συχνότητας του F1.

Η θεμελιώδης συχνότητα και η τροχιά της εκτιμάται μέσω μιας αυτοπαλινδρομητικής αναλύσεως (AutoRegressive - AR), η οποία είναι μια παραμετρική φασματική τεχνική και συνίσταται στην αλγοριθμική τεχνική μπλοκ δεδομένων που κάνει συνδυασμένη ελαχιστοποίηση του τετραγωνικού σφάλματος γραμμικής πρόβλεψης (Linear Prediction) τόσο προς τα εμπρός όσο και προς τα πίσω. Η τεχνική αυτή λέγεται τροποποιημένη συνδιακύμανση (modified covariance) και έχουν αναπτυχθεί γρήγοροι αλγόριθμοι για την λύση των κανονικών εξισώσεων της¹⁷. Κατόπιν διαφόρων δοκιμών καταλήξαμε στο μήκος των 256 δειγμάτων ανά μπλοκ και στην τάξη 45 για το φίλτρο AR και σχεδιάσαμε δισδιάστατα φάσματα (σχ. 3(b)) που χρησιμοποιήθηκαν για τον προσδιορισμό της τροχιάς της θεμελιώδους αρμονικής καθώς και τρισδιάστατα φάσματα (σχ. 4(a)) για την επιβεβαίωση των formants.

Η επίδραση του ισοκρατήματος στην εκτίμηση των formants εύκολα μπορεί να απομονωθεί εξ αιτίας των αμελητέας ισχύος υψηλών αρμονικών του σε σχέση με την ψαλμωδία, όσον αφορά στα υψηλά formants. Για τα χαμηλά formants και ειδικά για το F1 το ίσον διακρίνεται από την σταθερότητα στην συχνότητα και από το ότι η συχνότητα του F1 δεν είναι ακέραιο πολλαπλάσιο της συχνότητας του ίσου.

Τελικώς, πρέπει να σημειωθεί ότι επειδή για ωρισμένους ψάλτες παρατηρείται και πέμπτο formant στην υψηλή περιοχή, στις περιπτώσεις που αυτό ήταν ίσης ισχύος με το τέταρτο με την συχνότητά τους απόσταση μικρή θεωρήσαμε την μέση τιμή συχνότητας μεταξύ F4 και F5, ενώ στις περιπτώσεις που ήταν πολύ ασθενές απλά το αγνοήσαμε για λόγους ενιαίας περιγραφής των φασμάτων για όλους τους ψάλτες.

4. Αποτελέσματα και συζήτηση

4.1 Φωνήεντα ομιλίας και ψαλμωδίας

TABLE II.
Οι μέσες τιμές συχνότητων των τριών πρώτων formants για την ομιλία στην Ελληνική γλώσσα (Φακοτάκης, 1986)

vowel \ Formants	F1 (Hz)	F2 (Hz)	F3 (Hz)
/a/	520	1650	2550
/e/	480	1610	2480
/i/	410	2070	2540
/o/	470	1620	2460
/u/	410	1200	2450

Η ανάλυση των 930 ηχητικών τμημάτων έδωσε τα αποτελέσματα του πίνακα I για όλα τα formants για τα δύο πρώτα σετ δεδομένων. Συγκρίνοντας τις τιμές των συχνότητων του πίνακα I με εκείνες του πίνακα II, ο οποίος περιέχει αντίστοιχες τιμές για την ομιλία¹⁸ συμπεραίνουμε τα ακόλουθα:

1) Η μέση φορμαντική συχνότητα του F2 είναι συστηματικά μικρότερη στην ψαλμωδία από εκείνη της ομιλίας και για τα δύο σετ για

¹⁶ An approach to the study of human perception of non-stationary complex sound pitch, Ogorodnikova, E. A., Sensornye Sistemy 2, 304-311, 1988

¹⁷ Δές υποσημ. 14.

¹⁸ New method for automatic speech recognition, Fakotakis N., Ph. D. diss, University of Patras, 1986

κάθε φωνήεν από

τα /a, e, i, o, u/.

2) Οι μέσες τιμές της συχνότητας του F1 και για τα δύο σετ είναι συγκρίσιμες με εκείνες της ομιλίας με μία μικρή τάση να είναι μεγαλύτερες για την ψαλμωδία εκτός του φωνήεντος /u/.

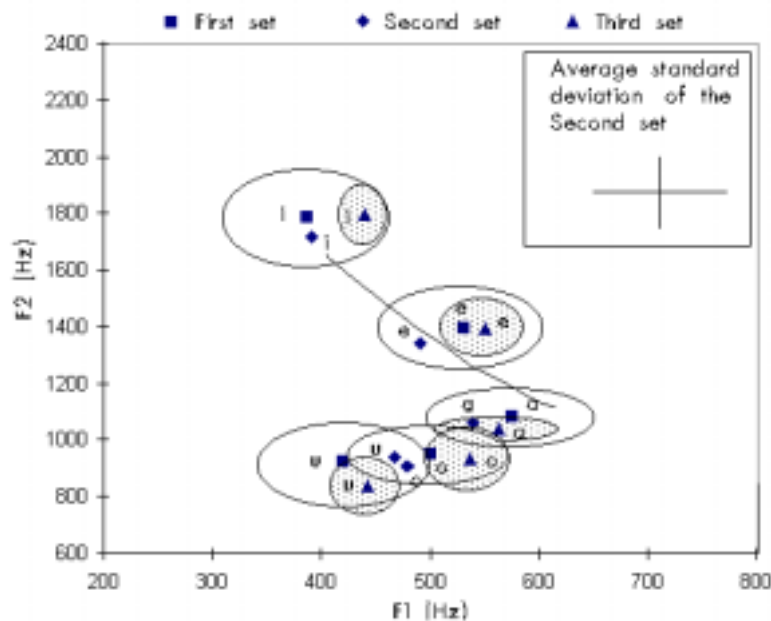
3) Η τρίτη φορμαντική συχνότητα είναι μικρότερη στην ψαλμωδία για τα έμπροσθεν φωνήεντα /e, i/ αλλά είναι μεγαλύτερη για τα όπισθεν /a, o, u/ και για τα δύο σετ (Σχ. 5).

TABLE III

Μέσες φορμαντικές συχνότητες για κάθε ψάλτη (από τους τρεις του τρίτου σετ) και φωνήεν, για τρεις ηχογραφήσεις σε τρεις διαφορετικές συνθήκες ηχογραφήσεως για τον καθένα.

Vowel \ Form	F1(Hz)	sdF1(Hz)	F2(Hz)	sdF2(Hz)	F3(Hz)	sdF3(Hz)	F4(Hz)	sdF4(Hz)	Third Set
/a/	487	36	1063	47	2560	233	3151	161	Psaltis1
/e/	523	16	1436	193	2349	33	3005	141	
/i/	414	29	1814	176	2256	128	2994	127	
/o/	563	77	991	208	2547	259	3161	190	
/u/	410	45	847	194	2723	98	3047	133	
/a/	563	43	1040	14	2656	84	3081	104	Psaltis2
/e/	511	42	1292	21	2416	31	3015	23	
/i/	420	17	1666	44	2295	65	3030	59	
/o/	480	27	862	82	2596	87	3064	51	
/u/	454	34	868	31	2785	120	3127	90	
/a/	634	81	1015	50	2852	113	3160	47	Psaltis3
/e/	614	60	1450	99	2651	133	3147	86	
/i/	484	13	1907	102	2543	131	3119	61	
/o/	563	26	948	56	2891	36	3133	104	
/u/	464	20	804	89	2889	71	3153	122	

Για να ερευνηθεί η επίδραση των μη ελεγχόμενων συνθηκών ηχογράφησης δημιουργήσαμε τα δεδομένα του τρίτου σετ. Μετρήθηκαν οι φορμαντικές συχνότητες για όλα φωνήεντα για κάθε ψάλτη (από τους τρεις) στις τρεις διαφορετικές συνθήκες. Πρώτα υπολογίστηκαν οι μέσες τιμές για κάθε φωνήεν και ψάλτη σε όλες τις συνθήκες καθώς και οι αποκλίσεις τους. Μετά βρέθηκαν οι μέσες τιμές για κάθε φωνήεν για όλους τους



Σχ. 6 Μέσες φορμαντικές συχνότητες των F1 και F2 για όλα τα σετ δεδομένων καθώς και οι τυπικές τους αποκλίσεις εμφανιζόμενες ως ημιάξονες των ελλείψεων. Στην πάνω δεξιά γωνία του σχήματος σχεδιάζονται οι τυπικές αποκλίσεις του δευτέρου σετ. Οι μέσες τιμές συχνότητων μεταξύ των τριών σετ είναι συγκρίσιμες όπως και οι τυπικές αποκλίσεις μεταξύ των δύο πρώτων σετ ενώ του τρίτου σετ είναι φανερά μικρότερες.

ψάλτες καθώς και οι αποκλίσεις τους. Ο πίνακας III φανερώνει ότι οι τυπικές αποκλίσεις για κάθε φωνήεν και φορμαντική συχνότητα είναι πολύ διαφορετικές (μέχρι και 10 φορές, φωνήεν /e/, F2) μεταξύ των ψαλτών, πράγμα που αποδεικνύει ότι οι διαφορές που οφείλονται στις ποικίλες συνθήκες είναι πολύ μικρότερες εκείνων που οφείλονται μεταξύ των ψαλτών. □Ετσι τα δεδομένα του πρώτου σετ δεδομένων μας οδηγούν σε ασφαλή συμπεράσματα.

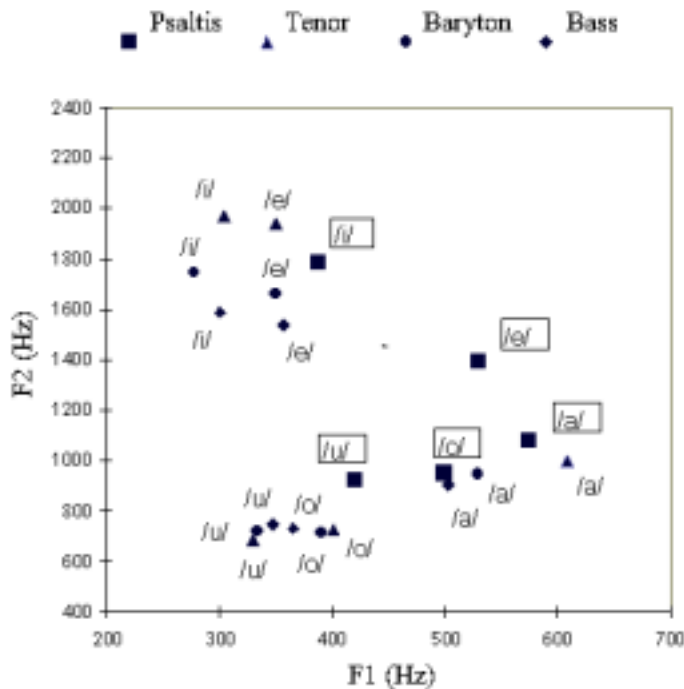
Σύγκριση μεταξύ των τριών σετ δεδομένων γίνεται στο σχήμα 6 στο οποίο για κάθε φωνήεν οι μέσες συχνότητες των formants είναι συγκρίσιμες ενώ οι περιοχές των σημείων για τα τρία σετ έχουν μια σημαντική επικάλυψη μεταξύ τους. Εδώ μπορούμε να δώσουμε μια πιθανή εξήγηση αυτών των αποτελεσμάτων. Είδαμε ότι η δεύτερη φορμαντική συχνότητα κυρίως εξαρτάται από την θέση της γλώσσας και η κίνησή της προς τον φάρυγγα δημιουργεί μια στένωση

που όσο μεγαλύτερη είναι τόσο ελαττώνεται η συχνότητα του F2. Η κίνηση προς τα πίσω της γλώσσας γίνεται μόνο στα όπισθεν φωνήεντα /a, o, u/ ενώ εδώ το F2 χαμηλώνει και στα έμπροσθεν φωνήεντα /e,

i/. Για τα πρώτα η στένωση είναι μεγαλύτερη στην ψαλμωδία ενώ στα δεύτερα πιο πιθανή εξήγηση είναι η επιμήκυνση της στοματικής κοιλότητας σε συνδυασμό με μια τάση να μοιάζουν όλα τα φωνήεντα προς τα όπισθεν φωνήεντα. Επίσης η ελαφρά ανύψωση του F1 μπορεί να εξηγηθεί με την φαρυγγική στένωση. Η ανύψωση του F3 για τα /o, u/ στην ψαλμωδία και η κάθοδός του για τα /e, i/ πρέπει να οφείλεται σε συνδυασμό της καθόδου του λάρυγγα, της επιμηκύνσεως της φωνητικής οδού και την θέση της γλώσσας ως προς την υπερώα.

4.2 Φωνήεντα ψαλμωδίας και όπερας.

Συγκρίνοντας τα δεδομένα φορμαντικών συχνοτήτων για τις τρεις κατηγορίες φωνών της όπερας¹⁹ του πίνακα



Σχ. 7 Πεδίο F1 - F2 για τα μέσα φωνήεντα τενόρων, βαρυτόνων, μπάσων και ψαλτών. Οι διαφορετικές θέσεις μεταξύ ψαλτών και οπερατικών τραγουδιστών υποδηλώνουν διαφορετική άρθρωση μεταξύ των δύο στύλ τραγουδιού.

TABLE IV
Μέσες φορμαντικές συχνότητες για Σουηδούς οπερατικούς τραγουδιστές: τενόρους, βαρυτόνους, και μπάσους μετρημένες από τον Thomas F. Cleveland (1977).

IV με εκείνα της ψαλμωδίας, για τα τρία πρώτα formants προκύπτουν τα εξής:

1) Σε όλα τα φωνήεντα το F1 formant είναι ψηλότερα στην ψαλμωδία, εκτός του φωνήεντος /a/ των τενόρων

2) Το formant F2 είναι ελαφρά υψηλότερο στην ψαλμωδία για τα όπισθεν φωνήεντα

3) Το F2 επίσης για το φωνήεν /i/ τοποθετείται, για την ψαλμωδία, στο πεδίο F1-F2 μεταξύ του αντιστοίχου των τενόρων και των βαρυτόνων αλλά είναι πέραν εκείνου των μπάσων (Σχ. 7).

Οι φορμαντικές συχνότητες ανάμεσα στην ομιλία και στο τραγούδι για τους οπερατικούς τραγουδιστές διαφέρουν²⁰ και υπάρχει διαφορετική άρθρωση μεταξύ τους. Ανάλογες διαφορές υπάρχουν και στην ψαλμωδία σε σχέση με την Ελληνική ομιλία.

Η ανύψωση του F1 στην ψαλμωδία μπορεί να εξηγηθεί αρθρωτικά με μια μεγαλύτερη στένωση του φάρυγγα από

εκείνη στην όπερα. Η ανύψωση του F2 για τα όπισθεν φωνήεντα στην ψαλμωδία φανερώνει πιθανώς την υψηλότερη θέση του λάρυγγα των ψαλτών. Επίσης η θέση του F2 για το φωνήεν /i/ μπορεί ίσως

να αποδοθεί σε μια ενδιάμεση εμπρόσθια γλωσσική στένωση μεταξύ τενόρων και βαρυτόνων. □Ομως τα ανωτέρω χρήζουν περαιτέρω αναλύσεως αρθρωτικών δεδομένων.

Μια συγκριτική μελέτη μεταξύ ψαλμωδίας και οπερατικού τραγουδιού για τα υψηλά formants μπορεί να γίνει βάσει του σχεσιογράμματος

¹⁹ Acoustic properties of voice timbre types and their influence on voice classification, Cleveland, T. F. JASA 61: 1622-29, 1977.

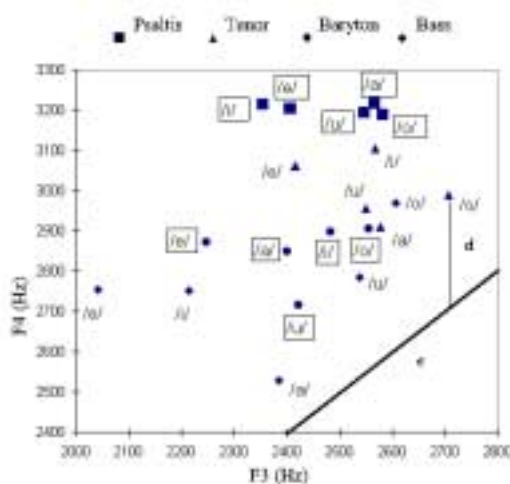
²⁰ Articulatory interpretation of the singing formant, Sundberg, J., JASA 55, 838-843, 1974

vowel \ Formant	F1 (Hz)	F2 (Hz)	F3 (Hz)	F4 (Hz)	
/a/	609	994	2576	2909	tenor
/e/	350	1942	2414	3061	
/i/	304	1969	2567	3105	
/o/	401	724	2706	2989	
/u/	330	682	2548	2957	bariton
/a/	530	944	2400	2849	
/e/	350	1662	2247	2873	
/i/	278	1744	2482	2897	
/o/	391	711	2554	2906	bass
/u/	333	719	2420	2716	
/a/	503	900	2386	2527	
/e/	356	1539	2041	2754	
/i/	300	1587	2214	2752	
/o/	365	729	2605	2969	
/u/	348	749	2536	2784	

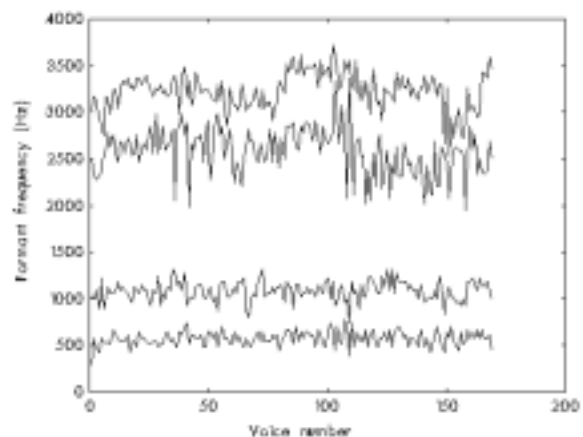
F3-F4 (Σχ. 8).
Παρατηρούμε τα παρακάτω:

- 1) Οι μέσες τιμές συχνοτήτων των F3 και F4 για τα έμπροσθεν φωνήεντα της ψαλμωδίας είναι πολύ κοντινές όπως επίσης και για τα όπισθεν φωνήεντα μεταξύ των.
- 2) Οι θέσεις των μέσων τιμών για την ψαλμωδία σημαντικά διαφέρουν από εκείνες για την όπερα.
- 3) Η διαφορά μεταξύ των φορμαντικών συχνοτήτων F3 και F4 παρουσιάζει μεγάλη ποικιλία για τις διάφορες ψαλτικές φωνές (Σχ. 9).

Είναι λογικό να θεωρήσουμε ότι η διακύμανση των μέσων φορμαντικών συχνοτήτων για ένα φωνήεν εξ αιτίας του διαφορετικού τονικού ύψους είναι πολύ μικρότερη σε σχέση με τους υπόλοιπους παράγοντες και επομένως μπορούμε να αγνοήσουμε την επίδραση των υψηλών pitch για την όπερα. Σύμφωνα με τον



Σχ. 8 Οι μέσες φορμαντικές συχνότητες των F3 και F4 για ψάλτες και οπερατικούς τραγουδιστές σε ένα σχεσιόγραμμα. Η κατακόρυφη απόσταση d ενός σημείου παριστάνει την διαφορά μεταξύ των συχνοτήτων των F4 και F3. Για τους ψάλτες εμφανίζεται η μεγαλύτερη απόσταση έναντι των οπερατικών τραγουδιστών.



Σχ. 9 Φορμαντικές συχνότητες για κάθε ψάλτη και το φωνήεν /a/. Είναι φανερό η διαφοροποίηση μεταξύ των ψαλτικών φωνών ως προς τα υψηλά formants.

J. Sundberg²¹ το F4 σχετίζεται άμεσα με το ηχόχρωμα και οφείλεται στις διαστάσεις του φωνητικού οργάνου. □Αρα λόγω της μεγάλης ποικιλίας φωνητικών οργάνων των ψαλτών πρέπει να υπάρχει μεγάλη διαφοροποίηση στην θέση του F4. Η θέση του F3 εξαρτάται μερικώς από τα φωνήεντα. □Ετσι η απόσταση μεταξύ F3 και F4 εξαρτάται από τα φωνήεντα. Τα F3, F4 και F5 formants συσπειρώνονται με την κάθοδο του λάρυγγα των οπερατικών τραγουδιστών και την ταυτόχρονη διεύρυνση του φάρυγγα σε περίπου 6 φορές το μέγεθος του λάρυγγα. Το αποτέλεσμα είναι ο σχηματισμός του singer's formant, το οποίο δημιουργεί μια προεξέχουσα κορυφή στο φάσμα γύρω στα 2800Hz περίπου και διαμορφώνει το ηχόχρωμα της φωνής σε πιο τραχύ. Στην ψαλμωδία αυτό δεν συμβαίνει παρόλο που τα υψηλά formants πλησιάζουν μεταξύ τους σε ωρισμένες φωνές διότι η απόστασή τους εξακολουθεί να παραμένει πολύ μεγαλύτερη εκείνης των οπερατικών φωνών. Επειδή η δυνατότητα μεταφοράς ήχου του

²¹ Βλ. ανωτέρω υποσημ. 8

φωνητικού οργάνου αυξάνεται κοντά και ανάμεσα σε δύο formants όταν η απόστασή τους μικρύνει²² προκύπτει ότι για την ψαλμωδία λόγω της μεγάλης σχετικά αποστάσεως μεταξύ των F3 και F4 η φωνή βρίσκεται πέραν του κρισίμου ορίου το οποίο καθορίζει την ανύψωση μιας νέας κορυφής στο φάσμα. Έτσι το ψαλτικό φωνητικό όργανο παρουσιάζει διαφορετική αρθρωτική σχηματοποίηση από το αντίστοιχο οπερατικό. Παράλληλα όταν εμφανίζονται περιπτώσεις ψαλτικών φωνών με κοντινά σε απόσταση τα F3 και F4 δεν συνδυάζονται με ένα τρίτο επίσης κοντινό formant για να συνενωθούν. Γι αυτόν τον λόγο το ψαλτικό ηχόχρωμα είναι διαφορετικό από το οπερατικό ενώ στα LTAS φάσματα εμφανίζονται δύο ξεχωριστές και μάλλον στενές κορυφές για την ψαλμωδία έναντι μιας ογκώδους κορυφής για την όπερα. Η κάθοδος του λάρυγγα υπάρχει πιθανότατα στην ψαλτική άρθρωση αλλά όχι με φαρυγγική διεύρυνση.

5. Συμπεράσματα

Σ' αυτήν την ανακοίνωση αναλύθηκε ένας μεγάλος αριθμός ψαλτικών - βυζαντινών φωνών, εκτιμήθηκαν οι φορμαντικές συχνότητες και συγκρίθηκαν μεταξύ τους και με εκείνες των οπερατικών τραγουδιστών. Τα σημαντικότερα ευρήματα ήσαν:

1) Οι ψάλτες άδουν κατά ένα διάφορο τρόπο από εκείνον όταν ομιλούν, πράγμα το οποίο μερικώς οφείλεται σε διαφορετική προφορά των φωνηέντων (ανύψωση του F1, κάθοδος του F2). Πιθανή εξήγηση βασίζεται στην επιμήκυνση του φωνητικού οργάνου, σε μια προς τα πίσω κίνηση της γλώσσας και ίσως στην κάθοδο της σιαγόνας.

2) Οι τραγουδιστάι της όπερας απαγγέλλουν διαφορετικά τα φωνήεντα από τους ψάλτες έχοντας την δευτέρα φορμαντική συχνότητα υψηλότερα από τους ψάλτες. Αρθρωτικά αυτό μπορεί να εξηγηθεί με μια μεγαλύτερη φαρυγγική στένωση του φωνητικού οργάνου των ψαλτών η οποία μπορεί να ενισχύεται με μια υψηλότερη θέση λάρυγγος έναντι των οπερατικών.

3) Η συχνотική απόσταση μεταξύ των υψηλών formants είναι πολύ μεγαλύτερη για την ψαλμωδία υποδηλώνοντας ένα διαφορετικό φωνητικό ηχόχρωμα για τους ψάλτες τόσο μεταξύ τους αλλά και εν συγκρίσει προς τους τραγουδιστάς της όπερας όπου το ηχόχρωμα είναι πολύ διαφορετικό με την εμφάνιση του singer's formant. Εμφανέστατα δεν παρατηρείται η ίδια τεχνική της καθόδου του λάρυγγα στους ψάλτες όπως στην όπερα. Τό F3 formant ξεχωρίζει τα όπισθεν από τα έμπροσθεν φωνήεντα υποδηλώνοντας ενιαία απαγγελία μεταξύ των ψαλτών.

ΕΥΧΑΡΙΣΤΙΕΣ

Οι συγγραφείς του παρόντος επιθυμούν να ευχαριστήσουν τους κ.κ. Νικ. Φακοτάκη (Πανεπιστήμιο Πατρών) για την διάθεση των δεδομένων Ελληνικής ομιλίας καθώς και Γεώργιο Κουρουπέτρογλου του τμήματος Πληροφορικής Παν. Αθηνών για την πολύτιμη βιβλιογραφία που μας παρείχε.

²² Βλ. υποσ. 8 επίσης

