

Ακουστική φωνητική μελέτη της παραγωγής και αναγνώρισης των φωνηέντων σε βαρήκοα άτομα

Α. Ε. ΒΑΚΑΛΟΣ

Δρ., Χειρουργός Ωτορινολαρυγγολόγος

Η διατριβή αυτή εκπονήθηκε υπό την επίβλεψη του καθηγητή Θ. Ηλιάδη, στο Εργαστήριο Ακοής και Ομιλίας της ΩΡΛ Κλινικής του ΑΠΘ στο Νοσοκομείο ΑΧΕΠΑ Θεσσαλονίκης, τα έτη 2003-2008.

Σκοπός

Ο σκοπός της διατριβής είναι διττός και περιλαμβάνει δύο σκέλη ανεξάρτητα μεταξύ τους:

- Μελέτη των αλλοιώσεων των ακουστικών χαρακτηριστικών των φωνηέντων και
- μελέτη των «λαθών» στην αναγνώριση μεμονωμένων φωνηέντων από βαρήκοα άτομα.

Ως αντικείμενο της εργασίας επιλέχθηκαν τα φωνήεντα (και όχι τα σύμφωνα ή κάποια άλλη συνιστώσα της προφορικής επικοινωνίας), διότι αυτά είναι ο «φορέας» και ο «πυρήνας» των συλλαβών, οι οποίες μερικές φορές αποτελούνται μόνο από ένα φωνήεν. Επίσης, χρησιμοποιούνται συχνότερα από τα σύμφωνα και μεταφέρουν περισσότερη ακουστική ενέργεια. Παράγονται, ενώ η φωνητική οδός είναι μάλλον ανοικτή, και μοναδική πηγή ήχου είναι το «βούισμα» που παράγεται από τη δόνηση των φωνητικών χορδών. Η ομάδα των φωνηέντων είναι επομένως πιο ομοιογενής όσον αφορά στα ακουστικά τους χαρακτηριστικά, τα οποία είναι απλά και σταθερά. Έχουν συνήθως μεγαλύτερη διάρκεια από τα σύμφωνα και είναι οι φορείς του «τονισμού» (προσωδίας) κατά την ομιλία. Επιπλέον, η βιβλιο-

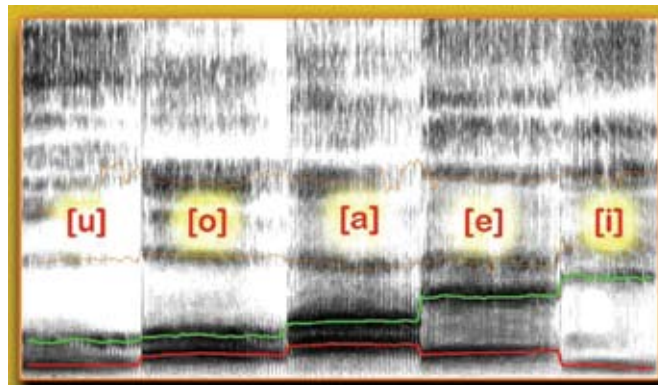
“ Ο σκοπός της διατριβής είναι διττός και περιλαμβάνει δύο σκέλη ανεξάρτητα μεταξύ τους: α) Μελέτη των αλλοιώσεων των ακουστικών χαρακτηριστικών των φωνηέντων και β) μελέτη των «λαθών» στην αναγνώριση μεμονωμένων φωνηέντων από βαρήκοα άτομα ”

γραφία είναι πλουσιότερη σε εργασίες που αφορούν στα φωνήεντα παρά στα σύμφωνα.

Βασικές έννοιες: Παραγωγή και αναγνώριση των φωνηέντων

Οι περισσότεροι φθόγγοι παράγονται μέσω της δόνησης των φωνητικών χορδών και λέγονται ηχηροί. Τέτοιοι φθόγγοι είναι όλα τα φωνήεντα και μερικά σύμφωνα (π.χ. [z], [ð], [b]). Τα υπόλοιπα σύμφωνα λέγονται άηχοι φθόγγοι και στην παραγωγή τους δε συμμετέχουν οι δονήσεις των φωνητικών χορδών (π.χ. [θ], [s], [t]).

Οι ηχηροί φθόγγοι παράγονται ως εξής:



Εικόνα 1. Φασματοτογράφημα των φωνηέντων της ελληνικής γλώσσας. Με κόκκινο φαίνεται ο F1, με πράσινο ο F2.

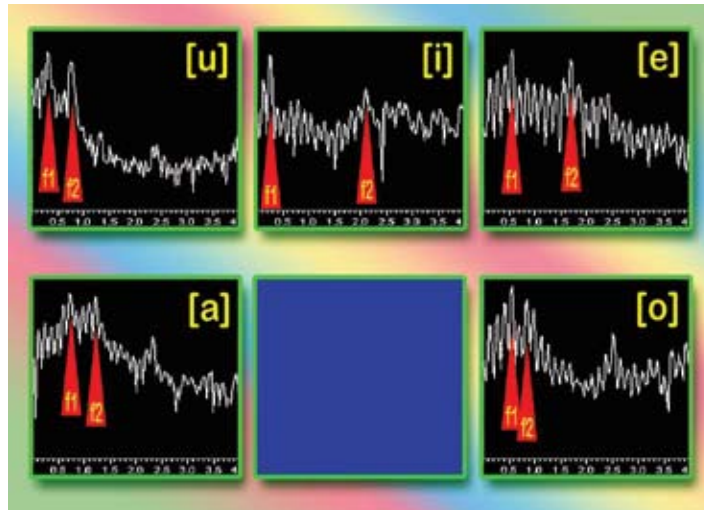
► Κατά τη φώνηση, η στήλη του εκπνεόμενου αέρα πιέζει από κάτω τις κλειστές φωνητικές χορδές. Καθώς η πίεση αυξάνει πέρα από κάποιο σημείο (το οποίο εξαρτάται από την τάση των φωνητικών χορδών και την πίεση του αέρα), η γλωττίδα υποχωρεί, ανοίγοντας στιγμιαία και επιτρέποντας την έξοδο μικρής ποσότητας αέρα με τη μορφή παλμού. Κατόπιν ξανακλείνει απότομα, λόγω της ελαστικότητας των ιστών, της μυϊκής τάσης και του φαινομένου Bernoulli, το οποίο προξενεί υποπίεση κατά την ταχεία δίοδο του αέρα από το στένωμα της γλωττίδας.

Έτσι, προκαλείται μια βραχεία δόνηση του αέρα που υπάρχει στο ανώτερο αναπνευστικό (φωνητικά όργανα), η οποία αποσβένεται σχεδόν αμέσως. Το φαινόμενο αυτό επαναλαμβάνεται περιοδικά, πολλές φορές ανά δευτερόλεπτο, και, λόγω της ταχείας απόσβεσης των παλμών, ο ήχος που προκύπτει (φωνή) μπορεί να αναλυθεί σε πλατύ φάσμα με θεμελιώδη συχνότητα ίση με το ρυθμό που ανοιγοκλείνει η γλωττίδα και πολλές αρμονικές συχνότητες (ακέραια πολλαπλάσια της θεμελιώδους).

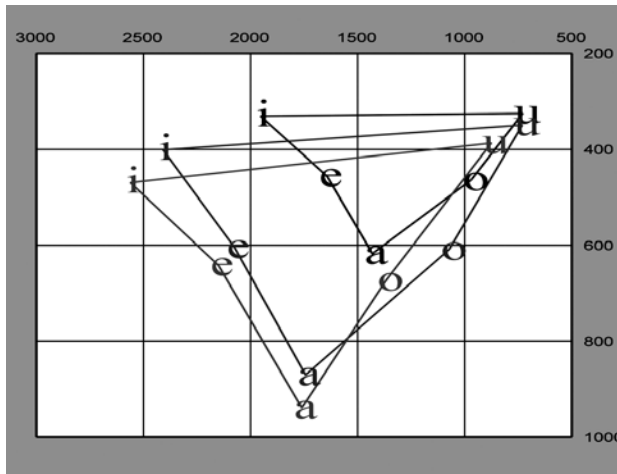
Φάσμα (spectrum) είναι η δυσδιάστατη γραφική απεικόνιση της σχετικής έντασης κάθε αρμονικής σε συνάρτηση με τη συχνότητα. Επειδή ο ήχος μεταβάλλεται συνεχώς, κάθε απεικόνιση του φάσματος αντιστοιχεί σε μια συγκεκριμένη στιγμή. Εάν προστεθεί και η συνιστώσα του χρόνου, έχουμε το φασματογράφημα (spectrogram), δηλαδή απεικόνιση των μεταβολών του φάσματος στο χρόνο. Μπορεί να απεικονισθεί και σε δύο διαστάσεις: Συνήθως, στον άξονα των x αντιστοιχεί ο χρόνος, των ψ η συχνότητα, ενώ η ένταση αποδίδεται χρωματικά.

Κατόπιν, τα φωνητικά όργανα (στόμα, φάρυγγας, γλώσσα, υπερώα), ανάλογα με το σχήμα που έχουν τη συγκεκριμένη στιγμή, λειτουργούν σαν αντηχείο που συντονίζεται και διαμορφώνει το φάσμα της φωνής σε φθόγγο. Δηλαδή, ενισχύουν κάποιες αρμονικές και εξασθενούν άλλες.

Μερικές από τις αρμονικές του φθόγγου έχουν ιδιαίτερα μεγάλη ένταση σε σχέση με τις υπόλοιπες και ονομάζονται διαμορφωτές (formants). Αυτοί είναι τα κύρια διακριτικά χαρακτηρι-



Εικόνα 2. Φάσμα των φωνηέντων της ελληνικής γλώσσας. Τα βέλη δείχνουν τους διαμορφωτές. Η κλίμακα είναι σε kHz.



Εικόνα 3. Χώρος των φωνηέντων της ελληνικής γλώσσας. Το μικρότερο τρίγωνο αντιστοιχεί στην παραγωγή των φωνηέντων από (φυσιολογικούς) άνδρες, το μεσαίο από γυναίκες και το μεγαλύτερο από παιδιά. Ο χώρος παραγωγής των φωνηέντων από γυναίκες είναι ευρύτερος και μετατοπισμένος προς τις ψηλές συχνότητες σε σχέση με αυτόν των ανδρών.

στικά των φωνηέντων. Συμβολίζονται με το γράμμα F και έναν αριθμό π.χ. F1: πρώτος διαμορφωτής, με τη χαμηλότερη συχνότητα -F2: δεύτερος διαμορφωτής, αυτός με την αμέσως ψηλότερη συχνότητα κ.ο.κ. Αντίστοιχα, η θεμελιώδης συχνότητα συμβολίζεται με το F0. Οι δύο πρώτοι διαμορφωτές είναι τα «κλειδιά» για την αναγνώριση των φωνηέντων, ιδιαίτερα της ελληνικής γλώσσας η οποία έχει 5 μόνο φωνήεντα: [a], [e], [u], [o] και [i] (εικόνες 1, 2).

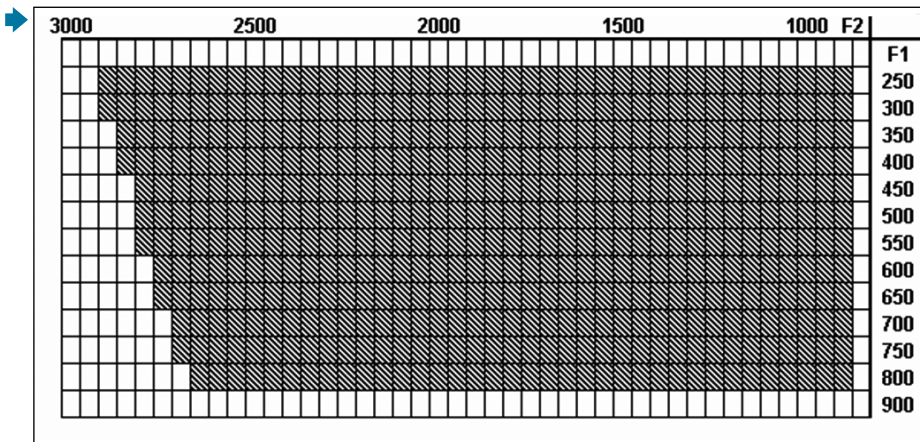
Τοποθετώντας το κάθε φωνήεν σε ένα δυσδιάστατο διάγραμμα, όπου στον άξονα των x -οριζόντιο- απεικονίζονται οι τιμές του F2 και στον άξονα των ψ -κάθετο- οι τιμές του F1, προκύπτει ένα σχήμα που ονομάζεται χώρος των φωνηέντων, το οποίο στην πραγματικότητα είναι ένα επίπεδο αδρά τρι-

γωνικού σχήματος (εικόνα 3).

Το σημείο που αντιστοιχεί στο κάθε φωνήεν αντιστοιχεί με τη θέση (κατά το οβελιαίο επίπεδο) του στόματος, στην οποία βρίσκεται η γλώσσα κατά την άρθρωση του φωνήεντος αυτού.

Υλικό

Συμμετείχαν 102 άτομα ηλικίας από 9 έως 81 ετών, από τα οποία τα 53 ήταν άνδρες. Τα 23 άτομα είχαν φυσιολογική ακοή και τα υπόλοιπα χωρίσθηκαν σε 7 ομάδες, ανάλογα με την αιτιολογία της βαρηκοΐας (πρεσβυακουσία-θορυβογενής-ωτοτοξική-συνδυασμός πρεσβυακουσίας με θορυβογενή-αιφνίδια βαρηκοΐα-βαρηκοΐα αντιλήψεως οφειλόμενη σε διάφορα ή πολλαπλά αίτια και βαρηκοΐα αγωγιμότητας ή μικτή). Επίσης, οι βαρήκοοι ►



Εικόνα 4. Τα σκιασμένα ορθογώνια είναι η περιοχή του χώρου των φωνηέντων που καταλαμβάνεται από τα 465 συνθετικά φωνήεντα.

χωρίστηκαν σε 6 κατηγορίες, ανάλογα με τη σοβαρότητα της πάθησής τους, οι οποίες αντιστοιχούν σε ελαφρού ως σοβαρού βαθμού βαρνηκοΐα.

Κανένα από τα άτομα που συμμετείχαν στη μελέτη δεν παρουσίαζε ακουστικά αντιληπτή αλλοίωση της φωνής και ομιλίας.

Η κατάταξη σε κατηγορίες, όσον αφορά στη μελέτη των χαρακτηριστικών της φωνής και τις δοκιμασίες ελεύθερου πεδίου, έγινε με βάση την ακοή στο καλύτερο αυτί.

Με αυτόν τον τρόπο κατάταξης, μόνο τρία άτομα έπασχαν από σοβαρού βαθμού βαρνηκοΐα, ενώ τα υπόλοιπα έπασχαν από βαρνηκοΐα μέτριου ως ελαφρού βαθμού.

Μέθοδος

Έγινε πλήρης κλινικός ΩΡΛ και ακολογικός έλεγχος σε όλους τους συμμετέχοντες, συμπεριλαμβανομένου του τονικού ακοογράμματος και της Ομιλητικής Δοκιμασίας Αναγνώρισης Ελληνικών Δισύλλαβων Λέξεων (η οποία είναι υπερουδική δοκιμασία διακριτικής ικανότητας και αποδίδει ποσοστό επιτυχούς αναγνώρισης λέξεων-WDS).

Από το σημείο αυτό και μετά, η εργασία χωρίστηκε στα δύο σκέλη.

Παραγωγή φωνηέντων

Μέθοδος και υπολογισμοί:

Ηχογραφήθηκε η φωνή όλων των εξεταζομένων κατά την εκφώνηση μιας φράσης που περιλάμβανε (πάνω

από μία φορά) και τα πέντε φωνήεντα της ελληνικής γλώσσας σε τονιζόμενες και άτονες συλλαβές και (δύο φορές) σειρά των ακόλουθων πέντε μονοσύλλαβων λέξεων: «πας», «πες», «πους», «πεις», «πως».

Κατόπιν, για κάθε εκφώνηση κάθε φωνηέντος από κάθε άτομο, βρέθηκε η θεμελιώδης (με το ελεύθερο πρόγραμμα «ανοικτού κώδικα» WaveSurfer 1.5.2 των Σουηδών Beskow και & Sjölander, τους οποίους και ευχαριστώ: <http://www.speech.kth.se/wavesurfer/>) και οι δύο πρώτοι διαμορφωτές και σχηματίστηκε ο χώρος παραγωγής των φωνηέντων από το συγκεκριμένο άτομο σε Hz και Bark.

Για κάθε εξεταζόμενο, καταχωρήθηκαν για όλες τις εκφωνήσεις κάθε φωνηέντος οι μέσες τιμές του F1 και F2 και οι λόγοι τυπικής απόκλισης προς μέση τιμή των F1 και F2 (σε Bark). Bark είναι μια ψυχοακουστική κλίμακα ύψους του ήχου, η οποία χωρίζει το ακουστικό φάσμα σε 24 ζώνες (δηλαδή από 1 έως και 24 Bark).

Κατόπιν έγιναν οι εξής υπολογισμοί, οι οποίοι αποδίδουν τα συνοπτικά χαρακτηριστικά του χώρου παραγωγής των φωνηέντων από κάθε άτομο:

- Μέσες τιμές F1 και F2 για όλα τα φωνήεντα. Απεικονίζουν τη θέση του «κέντρου βάρους» του χώρου φωνηέντων του εξεταζόμενου. Είναι δύο αριθμοί (σε Bark).
- Άθροισμα των λόγων τυπικής απόκλισης προς μέση τιμή των F1 και F2 για όλα τα φωνήεντα. Είναι ένας απόλυτος αριθμός, μέτρο της «δια-

σποράς» ή, αντίστροφα, της «σαφήνειας» του χώρου.

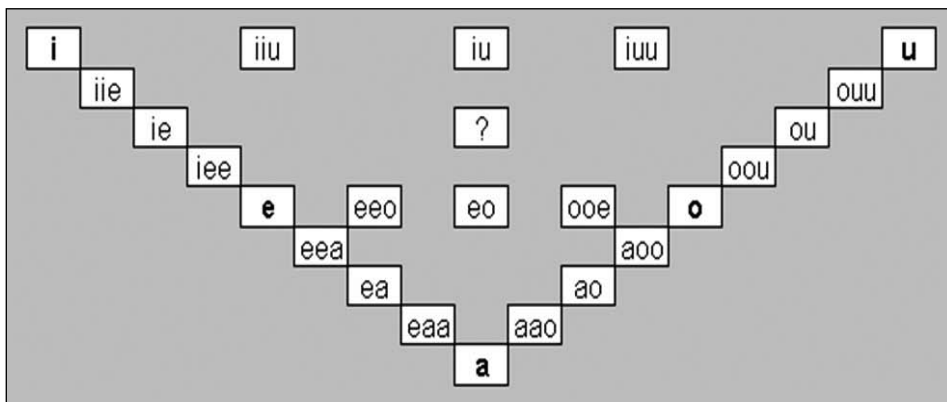
- Εμβαδό του χώρου παραγωγής των φωνηέντων του εξεταζόμενου, βασισμένο στις μέσες τιμές των F1 και F2. Είναι ένας αριθμός (σε τετραγωνικά Bark).

Τέλος, για κάθε ομάδα (αιτιολογίας ή βαθμού της βαρνηκοΐας) ως σύνολο και ξεχωριστά για άνδρες και γυναίκες έγιναν οι παρακάτω υπολογισμοί:

- Μέσες τιμές F1 και F2 των θέσεων «κέντρου βάρους» του χώρου παραγωγής φωνηέντων όλων των μελών της ομάδας. Δύο αριθμοί (σε Bark), οι οποίοι απεικονίζουν τη θέση «κέντρου βάρους» του χώρου φωνηέντων της ομάδας.
- Ευκλείδειο άθροισμα των παραπάνω μέσων τιμών (τετραγωνική ρίζα του αθροίσματος των τετραγώνων-Πυθαγόρειο θεώρημα). Είναι η απόσταση (σε Bark) του κέντρου του χώρου φωνηέντων της ομάδας από την αρχή των κλιμάκων.
- Μέση τιμή «διασποράς» των μελών της ομάδας. Απόλυτος αριθμός, μέτρο της διασποράς του χώρου της ομάδας.
- Μέση τιμή των εμβαδών του χώρου παραγωγής των φωνηέντων από όλα τα μέλη της ομάδας. Είναι ένας αριθμός (σε τετραγωνικά Bark).

Συμπεράσματα από την παραγωγή των φωνηέντων

Από τη στατιστική επεξεργασία των παραπάνω ευρημάτων προέκυψαν τα



Εικόνα 5. Οι θέσεις στις οποίες μπορούσε ο εξεταζόμενος να «τοποθετήσει» το συνθετικό φωνήεν. Είναι οι θέσεις των 5 πραγματικών φωνηέντων της νεοελληνικής γλώσσας, μία ακαθόριστη θέση (με το ερωτηματικό) και 18 ενδιάμεσες θέσεις.

ακόλουθα συμπεράσματα:

Ο χώρος παραγωγής φωνηέντων των γυναικών, σε σχέση με αυτόν των ανδρών, βρέθηκε μετατοπισμένος προς τις υψηλές συχνότητες, με μεγαλύτερη διασπορά και μεγαλύτερη έκταση.

Παρατηρήθηκαν ελάχιστες στατιστικά σημαντικές διαφορές κατά τη σύγκριση των παραπάνω παραμέτρων των διαφόρων «ομάδων αιτιολογίας» και «κατηγοριών βαθμού» των βαρικών με τις αντίστοιχες παραμέτρους των ατόμων με φυσιολογική ακοή.

Το τελικό συμπέρασμα είναι ότι τα άτομα με ελαφρού-μέτριου βαθμού βαρκοϊα δεν παρουσιάζουν σημαντικές αλλοιώσεις στην παραγωγή των φωνηέντων.

Αναγνώριση των φωνηέντων

Για τον έλεγχο των διαταραχών στην αναγνώριση των φωνηέντων, χρησιμοποιήθηκαν ήχοι που «κατασκευάστηκαν» τεχνητά και μοιάζουν με τα πραγματικά φωνήεντα της ελληνικής γλώσσας.

Δημιουργήθηκαν (το 1995, από τους Hawks και Φουράκη, οι οποίοι και μας τις παραχώρησαν για την παρούσα διατριβή) δύο ομάδες συνθετικών φωνηέντων, η μία με ανδρική θεμελιώδη συχνότητα (140Hz) και η άλλη με γυναικεία (240Hz). Οι κεντρικές συχνότητες των διαμορφωτών για κάθε συνθετικό φωνήεν καθορίστηκαν κατά «βήματα» των 50Hz. Από 250Hz ως 800Hz (12 βήματα) για τον πρώτο διαμορφωτή (F1) και από 900Hz έως 3000Hz (43 βήματα) για το δεύτερο διαμορ-

φωτή (F2). Ορισμένοι ήχοι απορρίφθηκαν και απέμειναν 465 συνθετικά φωνήεντα σε κάθε ομάδα (εικόνα 4). Κατόπιν, ακροατές με μητρική γλώσσα την ελληνική και φυσιολογική ακοή άκουσαν όλους αυτούς του 465+465 ήχους και επέλεξαν 151, τους οποίους θεώρησαν ότι μοιάζουν αρκετά με τα πραγματικά φωνήεντα της νεοελληνικής γλώσσας.

Στη δική μας εργασία, οι συμμετέχοντες άκουγαν αυτά τα 151 συνθετικά φωνήεντα και προσπαθούσαν να τα ταυτίσουν με κάποιο από τα πραγματικά φωνήεντα της ελληνικής γλώσσας ή (εφόσον αυτό το θεωρούσαν ανέφικτο) να τα «τοποθετήσουν» σε κάποια ενδιάμεση θέση του χώρου των φωνηέντων. Είχαν τη δυνατότητα να επιλέξουν κάποια από τις συνολικά 24 θέσεις (εικόνα 5). Για τις ανάγκες της εργασίας δημιουργήθηκε ειδική εφαρμογή βάσης δεδομένων Access - Visual Basic με την οποία εκτελούνταν όλες οι δοκιμασίες και καταχωρούνταν τα στοιχεία που προέκυπταν.

Η επιτυχία της προσπάθειας του εξεταζόμενου χαρακτηριζόταν από το πόσο κοντά βρισκόταν η θέση, στην οποία «τοποθετήθηκε» ο συνθετικός ήχος με τη θέση, στην οποία αυτός αντιστοιχούσε εκ κατασκευής. Όταν η θεμελιώδης του συνθετικού φωνήεντος ήταν γυναικεία, ως βάση των υπολογισμών χρησιμοποιήθηκε ο χώρος παραγωγής των φωνηέντων από γυναίκες, ενώ στην αντίθετη περίπτωση, από άνδρες.

Σα μέτρο σύγκρισης χρησιμοποιήθηκαν οι επιδόσεις των 20 καλύτερων από τα άτομα με φυσιολογική ακοή.

Συμπεράσματα από την αναγνώριση των συνθετικών φωνηέντων

Από τη σύγκριση των επιδόσεων στην αναγνώριση των συνθετικών φωνηέντων με άλλες παραμέτρους, βρέθηκαν οι παρακάτω συσχετίσεις:

- Οι επιδόσεις στην αναγνώριση των συνθετικών φωνηέντων συσχετίζονται ισχυρά με τις επιδόσεις στην ομιλητική ακοομετρία (δοκιμασία αναγνώρισης ελληνικών δισύλλαβων λέξεων).
- Αντιθέτως, υπάρχει ασθενής μόνο συσχέτιση των επιδόσεων στην αναγνώριση των συνθετικών φωνηέντων με τις επιδόσεις στο τονικό ακοόγραμμα και την κατηγορία βαθμού βαρκοϊας.
- Δε βρέθηκε αξιόλογη διαφοροποίηση των επιδόσεων μεταξύ των διαφόρων ομάδων αιτιολογίας της βαρκοϊας.
- Δε βρέθηκε συσχέτιση με την ηλικία.
- Αναγνωρίστηκαν ευκολότερα τα εμπρόσθια φωνήεντα [i] και [e] και δυσκολότερα το [u].

Η Δοκιμασία Εντοπισμού των Συνθετικών Φωνηέντων που αναπτύχθηκε για την παρούσα διατριβή είναι μια υπερουδική δοκιμασία που δεν είναι τονική, δεν είναι όμως ούτε ακριβώς ομιλητική. Είναι τελείως απαλλαγμένη από συνειρμό και, επομένως, βρίσκεται έξω από το γλωσσικό σύστημα επικοινωνίας. Με βάση αυτά τα χαρακτηριστικά, μπορεί να χρησιμοποιηθεί σε διάφορους τομείς. □