

# Εκπαιδευτικές εφαρμογές εργαλείων τεχνολογίας φωνής στη διδασκαλία της ξένης γλώσσας

Φ. Χαραλαμποπούλου\*, Δρ. Γ. Σταϊνχάουερ\*\*, Δρ. Σ. Μπακαμίδης\*\*,  
Καθ. Γ. Καραγιάννης\*\*\*

\*Υπ. Διδάκτωρ Παν/μιου Αθηνών-Ινστιτούτο Επεξεργασίας του Λόγου  
\*\*Ερευνητής-Ινστιτούτο Επεξεργασίας του Λόγου

\*\*\*Καθηγητής ΕΜΠ, Διευθυντής Ινστιτούτου Επεξεργασίας του Λόγου  
frieda@ilsp.gr, stein@ilsp.gr, bakam@ilsp.gr, gkara@ilsp.gr

## ΠΕΡΙΛΗΨΗ

Η διδασκαλία της ξένης γλώσσας αποτελεί ένα γνωστικό πεδίο στο οποίο οι Τεχνολογίες της Πληροφορίας και της Επικοινωνίας βρίσκουν πρόσφορο έδαφος για να αξιοποιηθούν οι δυνατότητές τους. Τόσο η κατάρτιση του γραπτού όσο και του προφορικού λόγου μπορεί πλέον να υποστηριχθεί από σύγχρονα υπολογιστικά εργαλεία και τεχνικές με αξιόπιστα αποτελέσματα. Σε σχέση με τον προφορικό λόγο και το φωνητικό-φωνολογικό επίπεδο της γλώσσας, αρχίζει πλέον να γίνεται δυνατή η χρησιμοποίηση εργαλείων Τεχνολογίας Φωνής (Speech Technology), που επιτρέπουν τον αυτόματο έλεγχο και την αξιολόγηση της προφοράς του ομιλητή με δυνατότητα ανατροφοδότησης (feedback), τον έλεγχο της ορθότητας της θέσης του τόνου, τη δυνατότητα πλοήγησης μέσα στο πρόγραμμα ή επίλυσης ασκήσεων με τη χρήση της φωνής του χρήστη αντί του ποντακιού κ.λπ.

Στην παρούσα εισήγηση, θα επιχειρηθεί μια αναφορά σε βασικές τεχνολογίες φωνής και σε μεθόδους εκπαιδευτικής αξιοποίησής τους στο πλαίσιο ηλεκτρονικών περιβαλλόντων μάθησης για τη διδασκαλία της ξένης γλώσσας γενικά, με επιμέρους αναφορές στη διδασκαλία της Ελληνικής ως ξένης γλώσσας, ειδικότερα.

**ΛΕΞΕΙΣ-ΚΛΕΙΔΙΑ:** Διδασκαλία ξένης γλώσσας, εργαλεία Τεχνολογίας Φωνής, αυτόματος έλεγχος προφορικού λόγου, αυτόματη αναγνώριση φωνής (Automatic Speech Recognition), σύνθεση φωνής (Speech Synthesis), αυτόματη φωνητική μεταγραφή λέξεων, εργαλείο ελέγχου της θέσης του τόνου, εργαλείο ελέγχου επιτονισμού και προσωδίας

## ΕΙΣΑΓΩΓΗ

Η επικοινωνιακή προσέγγιση στη διδασκαλία της ξένης γλώσσας –η οποία δεσπόζει τις τελευταίες δεκαετίες στο προσκήνιο της γλωσσοδιδασκτικής– μετατόπισε το ενδιαφέρον από την απόκτηση της γνώσης του δομολειτουργικού συστήματος της γλώσσας (γραμματική-συντακτικό), στην απόκτηση γνώσης όχι μόνο της ορθής αλλά και της κατάλληλης χρήσης της γλώσσας στις διάφορες περιστάσεις επικοινωνίας, με βάση πολιτισμικές και κοινωνικές νόρμες (δηλαδή την ταυτότητα των συνομιλητών, την εγγύτητα της σχέσης τους, τον χώρο στον οποίον διαμείβεται ο διάλογος κ.λπ.). Το πρωταρχικό ζητούμενο δεν είναι πλέον η γλωσσική επάρκεια, αλλά η επικοινωνιακή δεξιότητα.

Η διδασκαλία της γλώσσας γενικά και της ξένης γλώσσας ειδικότερα αποτελεί ένα γνωστικό πεδίο στο οποίο οι νέες τεχνολογίες βρίσκουν πρόσφορο έδαφος για να αξιοποιηθούν οι

πολύπλευρες δυνατότητές τους. Οι Τεχνολογίες της Επικοινωνίας και της Πληροφορίας μπορούν να συμβάλουν στη δημιουργία «αυθεντικών» επικοινωνιακών περιβαλλόντων, που προσφέρονται για την ανάπτυξη των επικοινωνιακών δεξιοτήτων και τη συνειδητοποίηση των μηχανισμών και λειτουργιών της υπό εκμάθηση γλώσσας. Η χρήση των πολυμεσικών στοιχείων προσφέρει πλουραλισμό ερεθισμάτων και μπορεί να καλύψει ένα ευρύ φάσμα μαθησιακών ιδιαιτεροτήτων. Η ενσωμάτωση ήχου, εικόνας, διαλόγων και κειμένων, βίντεο και κίνησης (animation), η δυνατότητα σύγχρονης και ασύγχρονης επικοινωνίας, καθώς και πρόσβασης σε αυθεντικά κείμενα στην υπό εκμάθηση γλώσσα (μέσω του Διαδικτύου) επιτρέπουν τη δημιουργία δυναμικών περιβαλλόντων ηλεκτρονικής μάθησης (e-learning), οδηγώντας την εκπαιδευτική και μαθησιακή πρακτική σε νέους δρόμους. Η σειριακή παρουσίαση των πληροφοριών καταργείται, στον διδασκόμενο παρέχεται πλέον η δυνατότητα ελεύθερης πλοήγησής του μέσα στην πληροφορία ανάλογα με τα εξατομικευμένα και διακριτά χαρακτηριστικά του, δηλαδή τις γλωσσικές του ανάγκες, την ηλικία του, τα ενδιαφέροντά του. Μετατρέπεται, έτσι, από παθητικό αποδέκτη γνώσης σε ενεργό συμμετοχο σε όλα τα στάδια της μάθησης, μέσα από ένα περιβάλλον συνεχούς αλληλεπίδρασης (interaction) και διερεύνησης.

Στο πλαίσιο τέτοιων επικοινωνιακών περιβαλλόντων μάθησης της ξένης γλώσσας, η παραγωγή τόσο του προφορικού όσο και του γραπτού λόγου μπορεί πλέον να υποστηριχθεί αυτόματα με την ενσωμάτωση σύγχρονων υπολογιστικών εργαλείων και τεχνικών με αξιόπιστα κατά κύριο λόγο αποτελέσματα.

Η κατάκτηση του φωνητικού και φωνολογικού συστήματος αποτελεί έναν συχνά παραμελημένο -και ίσως πολύ σημαντικό- τομέα στην εκμάθηση της ξένης γλώσσας. Ιδιαίτερα για τους ενήλικες, η παραγωγή των φωνημάτων και φθόγγων της ξένης γλώσσας με όσο το δυνατόν λιγότερο ξενική προφορά αποτελεί ένα ιδιαίτερα δύσκολο σημείο στη διαδικασία εκμάθησης. Σε σχέση με το φωνητικό-φωνολογικό επίπεδο, αρχίζει πλέον να καθίσταται δυνατή η χρησιμοποίηση εργαλείων Τεχνολογίας Φωνής (Speech Technology), που επιτρέπουν τον έλεγχο και την αξιολόγηση της προφοράς του ομιλητή με δυνατότητα ανατροφοδότησης (feedback), τον έλεγχο της ορθότητας της θέσης του τόνου, ή τη δυνατότητα πλοήγησής μέσα στο πρόγραμμα και επίλυσης ασκήσεων με τη χρήση της φωνής του χρήστη αντί του ποντικιού και του πληκτρολογίου.

Στόχος της παρούσας εισήγησης είναι η περιγραφή βασικών τεχνολογιών φωνής, καθώς και η διερεύνηση των δυνατοτήτων εκπαιδευτικής αξιοποίησής τους, με σκοπό την ανάπτυξη και βελτίωση γλωσσικών δεξιοτήτων που σχετίζονται με την παραγωγή του προφορικού λόγου και με το φωνητικό-φωνολογικό σύστημα της ξένης γλώσσας. Συγκεκριμένα, θα γίνει αναφορά στα εργαλεία αναγνώρισης και σύνθεσης φωνής, αυτόματης φωνητικής μεταγραφής, ελέγχου της θέσης του τόνου και του επιτονισμού, και στο πώς θα μπορούσαν τα εργαλεία αυτά να αξιοποιηθούν σε ηλεκτρονικά περιβάλλοντα μάθησης για την εκμάθηση της ξένης γλώσσας γενικά, με επιμέρους αναφορές στη διδασκαλία της Ελληνικής ως ξένης γλώσσας, ειδικότερα. Θα πρέπει εδώ να επισημανθεί ότι οι μέθοδοι αξιοποίησης των εργαλείων τεχνολογίας φωνής που προτείνονται εδώ αποτελούν ένα δείγμα των δυνατοτήτων πολύπλευρης αξιοποίησής τους.

## **ΕΡΓΑΛΕΙΑ ΤΕΧΝΟΛΟΓΙΑΣ ΦΩΝΗΣ ΚΑΙ ΕΚΠΑΙΔΕΥΤΙΚΗ ΑΞΙΟΠΟΙΗΣΗ ΤΟΥΣ**

Τις τελευταίες δύο δεκαετίες, η ανάπτυξη δεξιοτήτων που σχετίζονται με την παραγωγή προφορικού λόγου έχει αναδειχθεί σε σημαντική παράμετρο στη διαδικασία της εκμάθησης μιας ξένης γλώσσας. Τα προγράμματα σπουδών για την ξένη γλώσσα εστιάζουν στην παραγωγή προφορικού λόγου με στόχο την επικοινωνιακή επάρκεια.

Η δυνατότητα των μαθητών να συμμετέχουν σε διαλόγους που αναφέρονται σε πραγματικές περιστάσεις επικοινωνίας (situations) στο πλαίσιο της επικοινωνιακής προσέγγισης θεωρείται μια από τις σημαντικότερες παραμέτρους στην εκμάθηση της ξένης γλώσσας. Αυτή η μεταστροφή από τον γραπτό στον προφορικό λόγο δημιούργησε μια τάση για εκπαιδευτικά λογισμικά που προσφέρουν τη δυνατότητα για έλεγχο και αξιολόγηση του προφορικού λόγου. Η ορθή προφορά αποτελεί μια από τις συνιστώσες που προσδιορίζουν την επικοινωνιακή δεξιότητα του χρήστη (Celce-Murcia et al., 1996). Κατά συνέπεια, η ανάπτυξη αξιόπιστων συστημάτων στο πλαίσιο της Υποβοηθούμενης από Υπολογιστή Γλωσσικής Διδασκαλίας (CALL), που είναι σε θέση να αξιολογούν την προφορά των εκφωνούμενων από τον χρήστη λέξεων ή φράσεων και να δίνουν ανατροφοδότηση (feedback) σε πραγματικό χρόνο (real-time) είναι μια ολοένα αυξανόμενη ανάγκη.

Τα σύγχρονα εκπαιδευτικά λογισμικά που αναπτύσσονται με την τεχνολογία των πολυμέσων συνήθως παρέχουν στον χρήστη -μεταξύ άλλων- τη δυνατότητα ηχογράφησης. Σε τέτοια μαθησιακά περιβάλλοντα, ο χρήστης έχει τη δυνατότητα να ηχογραφεί τη φωνή του προφέροντας λέξεις ή φράσεις στην υπό εκμάθηση γλώσσα, να τη συγκρίνει με τη φωνή του φυσικού ομιλητή και να αξιολογεί ο ίδιος την προφορά του. Με την υποστήριξη τεχνολογιών φωνής, η αυτοαξιολόγηση του χρήστη μπορεί πια να πραγματοποιηθεί με οπτική ανατροφοδότηση, είτε με τη μορφή γραφικών αναπαραστάσεων του φωνητικού σήματος, είτε απεικονίζοντας τις κινήσεις του στόματος και των χειλιών, τη φωνητική οδό κ.ά.

Από διάφορες έρευνες και πειράματα που έχουν διεξαχθεί, έχει αποδειχτεί ότι οι γραφικές απεικονίσεις των χαρακτηριστικών του σήματος φωνής συμβάλλουν στη βελτίωση της προφοράς των λέξεων (Bernstein & Christian, 1996), αλλά και στον εν γένει ρυθμό (Markham & Nagano-Madesen, 1997) σε επίπεδο προτασιακού ή και κειμενικού λόγου. Ο συνδυασμός ηχητικής και οπτικής ανατροφοδότησης οδηγεί σε καλύτερες επιδόσεις. Με βάση τη διαπίστωση αυτή, η ηχογράφηση της φωνής του χρήστη θα ήταν χρήσιμο να δηλώνεται με γραφήματα που μορφοποιούν το φωνητικό σήμα, καθιστώντας το έτσι πιο απτό. Μια τέτοια γραφική απεικόνιση επιτρέπει τη δυνατότητα αυτοαξιολόγησης και σύγκρισης: για παράδειγμα, το γράφημα που παρήγαγε ο χρήστης μπορεί να αντιπαραβάλλεται με το αντίστοιχο γράφημα του φυσικού ομιλητή και ο πρώτος να προσπαθεί να ελαχιστοποιήσει την απόκλιση, πράγμα που ερμηνεύεται ως συνεχής προσπάθεια από πλευράς του για βελτίωση της προφοράς στην υπό εκμάθηση γλώσσα.

Ένα είδος γραφικής απεικόνισης του φωνητικού σήματος αποτελεί το ηχογράμμα (sonogram), το οποίο αποτελεί ειδική φασματική απεικόνιση του σήματος φωνής του ομιλητή σε συνάρτηση με το χρόνο. Άλλες γραφικές αναπαραστάσεις αφορούν στις καμπύλες χαρακτηριστικών της προσωδίας, όπως η περιβάλλουσα (contour) της θεμελιώδους συχνότητας (pitch) και οι μεταβολές των διαρκειών των στοιχειωδών ήχων και της έντασης του σήματος της φωνής. Εναλλακτικά, αντί για τη δυνατότητα ελέγχου και αξιολόγησης της σωστής εκφοράς των λέξεων μέσω οπτικής ανατροφοδότησης, η αξιολόγηση της προφοράς του χρήστη μπορεί να περιορίζεται σε απλή βαθμολογική ένδειξη -δεδομένου ότι ορισμένοι μπορούν να ισχυριστούν ότι η αποκωδικοποίηση και ερμηνεία τέτοιου είδους γραφικών αναπαραστάσεων και κυματομορφών από μη ειδικούς, είναι μια διαδικασία που προϋποθέτει μια περίοδο εξάσκησης και εξοικείωσης.

Ορισμένες βασικές τεχνολογίες φωνής που μπορούν να λειτουργήσουν συνδυαστικά και να παράσχουν τη δυνατότητα εκπαιδευτικής αξιοποίησης και ενσωμάτωσής τους σε εκπαιδευτικά λογισμικά για τη διδασκαλία της ξένης γλώσσας και ειδικότερα για την αξιολόγηση και βελτίωση της προφοράς, είναι οι ακόλουθες:

- Αυτόματη φωνητική μεταγραφική λέξεων

- Σύνθεση φωνής
- Αυτόματη αναγνώριση φωνής
- Εργαλείο ελέγχου τονισμού
- Εργαλείο ελέγχου επιτονισμού

### **Αυτόματη φωνητική μεταγραφή λέξεων**

Το εργαλείο αυτό επιτρέπει την αυτόματη φωνητική μεταγραφή οποιασδήποτε λέξης ή φράσης πληκτρολογήσει ο χρήστης, σύμφωνα με το Διεθνές Φωνητικό Αλφάβητο (International Phonetic Alphabet). Προκειμένου για μια γλώσσα όπως η Ελληνική ή η Ρωσική για παράδειγμα, η λειτουργία αυτή αποδεικνύεται ιδιαίτερα χρήσιμη, δεδομένου ότι το ελληνικό και το κυριλλικό αλφάβητο αντίστοιχα διαφοροποιείται πλήρως από το λατινικό και συνεπώς παρουσιάζει ιδιαίτερες δυσκολίες για μια μεγάλη μερίδα ξένων μαθητών, ιδιαίτερα στα πρώτα επίπεδα γλωσσομάθειας.

Το εργαλείο της αυτόματης φωνητικής μεταγραφής βασίζεται σε ένα σύνολο απλών αλλά και σύνθετων κανόνων που καλύπτουν το σύνολο της γλώσσας και μετατρέπουν το ορθογραφικό κείμενο -είτε αυτό είναι μία λέξη ή μία πρόταση- στο αντίστοιχο του φωνητικό. Σε περιπτώσεις όπου υπάρχουν διαφορετικές εκδοχές της προφοράς ορισμένων στοιχείων της γλώσσας από ομάδες πληθυσμού που ζουν σε συγκεκριμένες γεωγραφικές περιοχές (δηλαδή στην περίπτωση των διαλεκτικών αλλομορφών), υιοθετείται αυτό που ονομάζουμε standard εκδοχή.

Θα πρέπει εδώ να αναφερθεί ότι οι κανόνες μεταγραφής δεν είναι σε θέση να επιλύσουν προβλήματα αμφισβησίας που προκύπτουν σε επίπεδο πρότασης (για παράδειγμα, προκειμένου για την Ελληνική γλώσσα, το σύστημα αυτόματης φωνητικής μεταγραφής δεν μπορεί να κάνει τον διαχωρισμό μεταξύ του θηλυκού ουσιαστικού «άδεια» και του ουδέτερου επιθέτου στον πληθυντικό «άδεια»). Η επίλυση τέτοιων προβλημάτων αμφισβησίας ενδεχομένως μπορεί να επιτευχθεί μόνο αν γίνει μορφοσυντακτική ανάλυση του προτασιακού περιβάλλοντος που εμφανίζεται η λέξη, πράγμα όμως που έχει ως επακόλουθο την αύξηση της υπολογιστικής πολυπλοκότητας.

### **Σύνθεση φωνής**

Ένα σύγχρονο εκπαιδευτικό λογισμικό πολυμέσων για τη διδασκαλία της ξένης γλώσσας περιέχει εκφωνήσεις λέξεων, προτάσεων αλλά και κειμενικού λόγου από φυσικούς ομιλητές, επιτρέποντας έτσι τη συνεχή έκθεση του χρήστη στα ακούσματα της υπό εκμάθηση γλώσσας. Πέρα όμως από τη δυνατότητα εκφωνήσεων από φυσική ανθρώπινη φωνή, η σύγχρονη τεχνολογία παρέχει σήμερα συστήματα συνθετικής φωνής, φωνής δηλαδή που παράγεται από τον ίδιο τον υπολογιστή.

Ο συνθέτης φωνής, όπως λέγεται το εν λόγω εργαλείο, μπορεί να μετατρέψει ένα γραπτό κείμενο σε υψηλής καταληπτότητας συνθετική φωνή, ενώ μπορεί να χειρίζεται κατάλληλα αριθμούς, ημερομηνίες, ακρωνύμια, συντομογραφίες κ.λπ. Ένας αποδοτικός αλγόριθμος επιτονισμού προσθέτει την κατάλληλη μελωδία στηριζόμενος στον τονισμό, τα σημεία στίξης κ.λπ., προσδίδοντας έτσι φυσικότητα στην συνθετική φωνή.

Ένα εκπαιδευτικό λογισμικό που έχει ενσωματώσει μια τέτοια λειτουργία, παρέχει τη δυνατότητα στον χρήστη να προσλαμβάνει πληροφορίες προφοράς για οποιαδήποτε λέξη πληκτρολογήσει. Η τεχνολογία αυτή θα μπορούσε να θεωρηθεί ως ένα βήμα μετά την φωνητική μεταγραφή (που αναφέρθηκε παραπάνω) στην εκπαίδευση χρηστών που δεν έχουν εξοικειωθεί με το φωνητικό και φωνολογικό σύστημα της ξένης γλώσσας, ενώ διευκολύνει ιδιαίτερα άτομα που

αντιμετωπίζουν δυσκολίες στην όραση: ένα σύστημα που υποστηρίζει μια τέτοια τεχνολογία δεν απαιτεί από τον χρήστη να εστιάζει το βλέμμα του στην οθόνη του υπολογιστή του. Το κείμενο διαβάζεται από τον συνθέτη και ο χρήστης προσλαμβάνει την πληροφορία ακουστικά.

Το άκουσμα ενός συνθέτη τα παλαιότερα χρόνια μπορεί να ξένιζε, με την έννοια ότι η ποιότητα της παραγόμενης φωνής από τον υπολογιστή δεν διέθετε την φυσικότητα της ανθρώπινης φωνής. Η τεχνολογία όμως έχει και εδώ προχωρήσει σημαντικά και η χροιά της φωνής που παράγει ένας σύγχρονος συνθέτης έχει πλέον βελτιωθεί σημαντικά και συνεχίζει συνεχώς να βελτιώνεται σε μια προσπάθεια να καταστεί πιο ελκυστική και να πλησιάσει –στο μέτρο του δυνατού- τη φυσική φωνή.

### **Αυτόματη αναγνώριση φωνής**

Η επεξεργασία και αναγνώριση του προφορικού λόγου στον άνθρωπο και στη μηχανή πραγματώνεται με τελείως διαφορετικό τρόπο (Bernstein et al., 1996). Στην περίπτωση του ανθρώπου, απαιτούνται πολύπλοκες νοητικές διεργασίες και επεξεργασίες προκειμένου να συσχετιστούν τα ακουστικά σήματα με τις λέξεις και τις σημασίες τους. Αντιθέτως, ο υπολογιστής «αντιλαμβάνεται» τον φωνούμενο λόγο ως μια σειρά ψηφιακών τιμών. Παρόλες τις θεμελιώδεις διαφορές, εν τούτοις και κατ' ουσίαν το πρόβλημα της αναγνώρισης και στις δύο περιπτώσεις εστιάζεται στο ίδιο πράγμα: στην αναζήτηση του κατάλληλου ταιριάσματος μεταξύ ενός συγκεκριμένου ήχου και της αντίστοιχης λεκτικής οντότητας. Η αυτόματη αναγνώριση φωνής επιχειρεί να βελτιστοποιήσει και να εξομοιώσει αυτού του είδους τις διεργασίες υπολογιστικά.

Στην αναγνώριση φωνής έχουν χρησιμοποιηθεί διάφορες προσεγγίσεις, όπως η Αναγνώριση Μεμονωμένων Λέξεων (Isolated Word Recognition), η Αναγνώριση Συνεχούς Λόγου (Continuous Speech Recognition) κ.ά. Τα συστήματα της πρώτης κατηγορίας αναγνωρίζουν ολόκληρες λέξεις ή φράσεις -οι οποίες όμως εκλαμβάνονται από το σύστημα ως μία οντότητα-, τις οποίες έχουν «μάθει» σε ένα αρχικό στάδιο, το στάδιο της εκμάθησης. Τα συστήματα αυτά αποτελούν τα πρώτα συστήματα αναγνώρισης φωνής. Στο πλαίσιο αυτών των συστημάτων, για την επίτευξη της αναγνώρισης ο ομιλητής έπρεπε να εκφωνεί κάθε λέξη ή φράση μεμονωμένα, δηλαδή μεταξύ των εκφωνούμενων λέξεων ή φράσεων ήταν απαραίτητο να παρεμβάλλονται παύσεις. Στα συστήματα Αναγνώρισης Συνεχούς Λόγου δεν υφίστανται τέτοιοι περιορισμοί και ο ομιλητής μπορεί πια να εκφωνεί οποιαδήποτε λέξη ή φράση με φυσικό ρυθμό. Κατά κανόνα, το σύστημα Αναγνώρισης Συνεχούς Λόγου στο στάδιο της εκμάθησης μαθαίνει μικρότερης διάρκειας ηχητικές μονάδες (π.χ. διφωνήματα, συλλαβές κ.ά.), βάσει των οποίων είναι στη συνέχεια σε θέση να αναγνωρίζει ολόκληρες λέξεις ή φράσεις.

Σε ένα λογισμικό για τη διδασκαλία της ξένης γλώσσας, η τεχνολογία της αναγνώρισης φωνής σε συνδυασμό με άλλα συστήματα ελέγχου της προφοράς -όπως για παράδειγμα τα εργαλεία επιτονισμού και ελέγχου του τόνου- μπορεί να έχει τις εξής εκπαιδευτικές εφαρμογές:

(i) *Δυνατότητα προφορικών απαντήσεων:* Στο πλαίσιο της υποβοηθούμενης από υπολογιστή γλωσσικής διδασκαλίας, η πάγια τακτική στη συμπλήρωση των απαντήσεων στηρίζεται στη χρήση του πληκτρολογίου ή του ποντικιού. Η αναγνώριση φωνής παρέχει πλέον τη δυνατότητα και προφορικών απαντήσεων. Σε μια τέτοια εφαρμογή, ο χρήστης καλείται αντί να δηλώσει την απάντηση με το ποντίκι ή το πληκτρολόγιο, απλά να την εκφωνήσει, παρακινούμενος έτσι να χρησιμοποιήσει την υπό εκμάθηση γλώσσα.

Οι ασκήσεις που προσφέρονται για μια τέτοιου είδους γλωσσική δραστηριότητα είναι κατά κύριο λόγο ασκήσεις του τύπου πολλαπλής επιλογής (multiple choice). Οι απαντήσεις που πρέπει να εκφωνήσει ο χρήστης σε τέτοιου είδους ασκήσεις μπορούν είτε να περιοριστούν σε επίπεδο μεμονωμένων λέξεων (και μάλιστα λέξεων που ανήκουν σε ένα κλειστό σύνολο, για παράδειγμα

απαντήσεις του τύπου «Σωστό–Λάθος», «Ναι-Όχι»), είτε να είναι ολόκληρες φράσεις ή πλήρεις προτάσεις. Οι συνεχείς εξελίξεις της τεχνολογίας επιτρέπουν πλέον στα συστήματα αναγνώρισης φωνής να χειρίζονται και να αναγνωρίζουν με επιτυχία τόσο μεμονωμένες λέξεις, όσο και προτασιακό λόγο.

(ii) *Λεξιλογικές ασκήσεις*: Η αναγνώριση φωνής μπορεί να αξιοποιηθεί σε σχέση με την εξάσκηση και βελτίωση της προφοράς σε ασκήσεις λεξιλογίου. Σε τέτοιου είδους ασκήσεις, ο χρήστης καλείται να εκφωνήσει το όνομα του συγκεκριμένου αντικείμενου που του ζητείται κάθε φορά. Σε περιβάλλον εκπαιδευτικών εφαρμογών για γλωσσική διδασκαλία σε παιδιά, για παράδειγμα, η εκφώνηση από τον ομιλητή-χρήστη και η αναγνώριση από το σύστημα θα μπορούσαν να έχουν τη μορφή ερεθίσματος-αντίδρασης. Συγκεκριμένα, η προφορά μιας λέξης που δηλώνει ένα αντικείμενο στην οθόνη, μπορεί να ενεργοποιεί το αντίστοιχο αντικείμενο με τη χρήση τεχνικών κίνησης (animation). Για παράδειγμα, το παιδί βλέπει στην οθόνη του υπολογιστή του μια εικόνα ενός ποδηλάτου, και καλείται να προφέρει τη λέξη «ποδήλατο». Εφόσον η λέξη προφερθεί σωστά και αναγνωριστεί από το σύστημα, το ποδήλατο ενεργοποιείται και αρχίζει να κινείται, ενώ στην αντίθετη περίπτωση μη σωστής εκφοράς/προφοράς της λέξης, το αντικείμενο παραμένει ανενεργό.

(iii) *Δυνατότητα πλοήγησης*: Η αναγνώριση φωνής μπορεί να αξιοποιηθεί για την υποβοήθηση της γενικότερης πλοήγησης μέσα στο πρόγραμμα, απελευθερώνοντας έτσι τον χρήστη από την ανάγκη συχνής χρήσης του ποντικιού για επαναλαμβανόμενες τυποποιημένες κινήσεις και εντολές, του τύπου: «πήγαινε στην επόμενη σελίδα», «πήγαινε στην προηγούμενη σελίδα», «άνοιξε το λεξικό», «έξοδος από το πρόγραμμα», «επιστροφή στην αρχική οθόνη» κ.τ.λ. Ο χρήστης, με άλλα λόγια, μπορεί να ηχογραφήσει τη φωνή του εκφωνώντας τη συγκεκριμένη εντολή στην υπό εκμάθηση γλώσσα, το σύστημα να την αναγνωρίζει, και η εντολή να εκτελείται. Είναι και αυτός ένας έμμεσος τρόπος προτροπής του χρήστη για τη χρήση του γλωσσικού κώδικα, προκειμένου να επιτευχθεί μια άλλου είδους επικοινωνία και αλληλεπίδραση, αυτή τη φορά μεταξύ χρήστη και μηχανής. Ο χρήστης εκφωνεί τις εντολές, το σύστημα «καταλαβαίνει» και ανταποκρίνεται, επικοινωνεί δηλαδή υπό μιαν έννοια με τον χρήστη. Σε περίπτωση μη αναγνώρισης της προφορικής εντολής -για την οποία χρησιμοποιείται ο τεχνικός όρος *απόρριψη* (rejection)- το σύστημα μπορεί να επιστρέφει στον χρήστη ένα είδος προκατασκευασμένης απάντησης του τύπου «Δεν κατάλαβα», «Παρακαλώ το επαναλαμβάνετε;» ή κάτι ανάλογο, προσομοιάζοντας έτσι μια πραγματική επικοινωνιακή κατάσταση, κατά την οποία ο δέκτης δεν αντιλαμβάνεται το μήνυμα του πομπού και ζητά την επανάληψή του.

Οι παραπάνω εφαρμογές της αναγνώρισης φωνής φαίνεται να είναι καταλυτικής σημασίας και να έχουν ιδιαίτερα χρηστικό χαρακτήρα για άτομα με ειδικές ανάγκες, και πιο συγκεκριμένα για άτομα που παρουσιάζουν κινησιακές δυσκολίες, οι οποίες δεν τους επιτρέπουν να κάνουν χρήση του ποντικιού και του πληκτρολογίου. Η τεχνολογία απελευθερώνει τους χρήστες αυτής της πληθυσμιακής ομάδας από τέτοιου είδους περιορισμούς και καθιστά δυνατή την επαφή τους με μαθησιακά περιβάλλοντα, που υποστηρίζουν τη δυνατότητα αλληλεπίδρασης μεταξύ χρήστη και μηχανής με βάση τον φωνομένο λόγο.

### **Εργαλείο ελέγχου της θέσης του τόνου**

Ο τόνος ανήκει στα προσωδιακά χαρακτηριστικά της φωνής. Ιδιαίτερα για την Ελληνική γλώσσα, ο τόνος συνιστά σημαντικό στοιχείο του φωνητικού-φωνολογικού επιπέδου της γλώσσας, δηλώνεται σαφώς στον γραπτό λόγο με το τονικό σημάδι, και λειτουργεί ως φορέας σημασίας: αλλαγή στη θέση του τόνου μιας λέξης συνεπάγεται πλήρως διαφοροποιημένες σημασιολογικά λέξεις (*γένια-γενιά, πότε-ποτέ* κ.τ.λ.). Κατά συνέπεια, λανθασμένη χρήση του τόνου μπορεί να οδηγήσει σε προβλήματα επικοινωνίας.

Μια από τις συνηθέστερες και απλούστερες πρακτικές που υιοθετείται στην πλειοψηφία των εκπαιδευτικών εφαρμογών για την εκμάθηση του τονισμού, είναι η ακόλουθη: εμφανίζονται στην οθόνη άτονες λέξεις και ο χρήστης καλείται να τοποθετήσει το τονικό σημάδι, ακούγοντας τις λέξεις να προφέρονται από φυσικό ομιλητή. Μια άλλη προσέγγιση είναι η γραφική απεικόνιση ορισμένων παραμέτρων σχετιζόμενων με την προσωδία και ειδικότερα με τον τόνο, η οποία όμως προϋποθέτει μια περίοδο εξοικείωσης με τέτοιου είδους κυματομορφές, για την πρόσληψη και σωστή ερμηνεία της παρεχόμενης πληροφορίας.

Οι τεχνολογικές εξελίξεις παρέχουν πλέον τη δυνατότητα αυτόματου ελέγχου του σωστού τονισμού της λέξης, με τη χρήση κατάλληλων εργαλείων. Η αξιοποίηση εργαλείων ελέγχου της θέσης του τόνου κρίνεται ιδιαίτερα σημαντική για την εκμάθηση του τόνου, που συνιστά μια από τις παραμέτρους της ορθής προφοράς. Τεχνικές ψηφιακής επεξεργασίας σήματος (digital signal processing) και αναγνώρισης προτύπων (pattern recognition) μπορούν να χρησιμοποιηθούν για τον αυτόματο προσδιορισμό της θέσης του τόνου σε προφερόμενες λέξεις, ενώ καταβάλλεται συνεχής ερευνητική προσπάθεια για περαιτέρω βελτίωση της αξιοπιστίας των συστημάτων αυτών.

Στο πλαίσιο των δυνατοτήτων που παρέχει σήμερα η τεχνολογία, μια πιθανή εκπαιδευτική αξιοποίηση του εργαλείου ελέγχου της θέσης του τόνου είναι η ακόλουθη: αρχικά, προτείνεται στον χρήστη μια λέξη, την οποία καλείται να προφέρει. Στη συνέχεια, το σύστημα αναγνωρίζει τη λέξη (μέσω του εργαλείου αυτόματης αναγνώρισης φωνής), εντοπίζει την τονιζόμενη συλλαβή και εμφανίζει γραπτώς τη λέξη, τοποθετώντας το τονικό σημάδι στη συλλαβή που εκφωνήθηκε από τον χρήστη. Ο χρήστης διευκολύνεται έτσι στο να εντοπίσει σε ποια συλλαβή τόνισε την λέξη και να επισημάνει τυχόν λάθη.

### **Εργαλείο ελέγχου επιτονισμού**

Η διδασκαλία της ορθής προφοράς σε μια ξένη γλώσσα δεν μπορεί ασφαλώς να περιορίζεται μόνο σε επίπεδο φωνημάτων (segmental/phonemic level), αλλά αφορά και στον επιτονισμό και την προσωδία γενικότερα, δηλαδή σε υπερκείμενα του φωνήματος επίπεδα (suprasegmental level), με χαρακτηριστικά που επεκτείνονται σε επίπεδο φράσης, πρότασης και τελικά συνεχούς λόγου (discourse). Η σωστή χρήση μάλιστα αυτών των υπερκείμενων του φωνήματος στοιχείων έχει αποδειχτεί ότι μπορεί να βελτιώσει τη συντακτική και σημασιολογική καταληπτότητα του προφορικού λόγου (Crystal, 1981).

Στην προφορική συνομιλία, αυτά τα χαρακτηριστικά διευκολύνουν τον ακροατή στο να εντοπίσει το όριο των προτάσεων, αλλά και να διαχωρίσει την καταφατική από την ερωτηματική πρόταση, ιδιαίτερα σε γλώσσες όπως τα Ελληνικά, όπου η διαφοροποίηση της κατάφασης από την ερώτηση πραγματώνεται κατά κύριο λόγο με την μεταβολή του επιτονισμού (π.χ. *Ο Γιάννης ήρθε το πρωί.* - *Ο Γιάννης ήρθε το πρωί.;*), σε αντίθεση με άλλες γλώσσες (όπως π.χ. τα αγγλικά), όπου η διαφοροποίηση αυτή δηλώνεται από αλλαγές στο μορφοσυντακτικό επίπεδο (π.χ. *John came in the morning-Did John come in the morning?*)

Τα τελευταία χρόνια διεξάγεται συστηματική έρευνα για την ανάπτυξη εργαλείων για τον έλεγχο του επιτονισμού στη διαδικασία εκμάθησης της ελληνικής γλώσσας. Όπως και στο προηγούμενο εργαλείο, η ανάπτυξη τέτοιων συστημάτων μπορεί να βασισθεί σε τεχνικές ψηφιακής επεξεργασίας σήματος και αναγνώρισης προτύπων, όπως η Δυναμική Χρονική Σύμπτωση (Dynamic Time Warping). Μια πιθανή εκπαιδευτική αξιοποίησή τους περιλαμβάνει την οπτικοποίηση μιας από τις συνιστώσες της προσωδίας, της καμπύλης θεμελιώδους συχνότητας (fundamental frequency curve) σε συνάρτηση με τον χρόνο. Ο χρήστης μπορεί να συγκρίνει οπτικά τις καμπύλες θεμελιώδους συχνότητας που παράγονται αυτόματα από το

σύστημα καθώς διαβάζει φράσεις ή προτάσεις, με τις αντίστοιχες καμπύλες που παρήχθησαν κατά την εκφώνηση των ίδιων φράσεων ή λέξεων από φυσικό ομιλητή και έχουν εκ των προτέρων καταχωρηθεί στο σύστημα. Σε πιο προηγμένα συστήματα ένα τέτοιο εργαλείο μπορεί επιπρόσθετα να επισημαίνει στον χρήστη τις περιοχές στις οποίες παρατηρείται μεγάλη απόκλιση.

Θα πρέπει στο σημείο αυτό να τονιστεί ότι η αυτόματη αξιολόγηση της προφοράς αποτελεί μια πολυσύνθετη διαδικασία και η διασφάλιση υψηλών ποσοστών επιτυχίας και αξιοπιστίας προϋποθέτει τη συνδυαστική εφαρμογή των εργαλείων που προαναφέρθηκαν.

## ΕΠΙΛΟΓΟΣ

Η επιστημονική κοινότητα που αντιμετωπίζει επιφυλακτικά ή και επικριτικά το όλο θέμα της εισβολής των Νέων Τεχνολογιών στην εκπαίδευση γενικά, και στην γλωσσική διδασκαλία ειδικότερα, επιχειρηματολογεί μιλώντας –μεταξύ πολλών άλλων- για απουσία ενός κοινού θεωρητικού πλαισίου για τον σχεδιασμό και την αξιολόγηση εκπαιδευτικών λογισμικών (Chapelle, 1997), για απουσία αποδεδειγμένης πρακτικής αξίας των παιδαγωγικών πλεονεκτημάτων της χρήσης των υπολογιστών στη μαθησιακή διαδικασία (Salaberry, 1996), αλλά και για τους περιορισμούς της ίδιας της τεχνολογίας και των δυνατοτήτων της (Holland, 1995 και Warschauer, 1996).

Σχετικά με τις τεχνολογίες φωνής ειδικότερα, ο προβληματισμός εστιάζεται στο κατά πόσον η διαθέσιμη τεχνολογία μπορεί πράγματι να ενσωματωθεί με αξιόπιστα αποτελέσματα στο πλαίσιο της διδασκαλίας υποβοηθούμενης από υπολογιστή, και όντως να συμβάλει στην ανάπτυξη και βελτίωση των γλωσσικών δεξιοτήτων που σχετίζονται με την παραγωγή του προφορικού λόγου. Το ενδιαφέρον εστιάζεται στην κατάλληλη χρήση και αξιοποίηση της τεχνολογίας αναγνώρισης φωνής σε συνδυασμό με άλλα εργαλεία ελέγχου της προφοράς.

Είναι γεγονός ότι παλαιότερα η τεχνολογία των συστημάτων αναγνώρισης φωνής επέβαλλε κάποιους περιορισμούς. Τα συστήματα αυτά λειτουργούσαν με μεγαλύτερα ποσοστά επιτυχίας όταν λαμβάνονταν υπόψη μια σειρά από συνιστώσες: αναγνώριση λέξεων που απέχουν φωνητικά μεταξύ τους, αναγνώριση σε επίπεδο μεμονωμένων λέξεων και όχι συνεχούς λόγου, έλλειψη θορύβου, η χρήση κατάλληλου και καλής ποιότητας μικροφώνου για την ηχογράφηση της φωνής του χρήστη κ.ά. Πολλά από τα τεχνικά προβλήματα του παρελθόντος έχουν ήδη ξεπεραστεί και η ποιότητα των αποτελεσμάτων συνεχώς βελτιώνεται. Στην τεχνολογία φωνής, η έρευνα σήμερα εστιάζεται στην ανάπτυξη συστημάτων αναγνώρισης φωνής που θα λειτουργούν με επιτυχία τόσο σε επίπεδο μεμονωμένων λέξεων, όσο και σε επίπεδο φράσεων ή και συνεχούς λόγου, ακόμα και σε θορυβώδη περιβάλλοντα, ή σε περιβάλλοντα όπου μιλούν ταυτόχρονα πολλοί ομιλητές (το λεγόμενο ‘cocktail party effect’).

Τα τελευταία χρόνια έχει ξεκινήσει μια ερευνητική προσπάθεια για την ανάπτυξη συστημάτων κατανόησης του λόγου (Speech Understanding), δηλαδή συστημάτων που όχι μόνο αναγνωρίζουν την ανθρώπινη φωνή αλλά είναι σε θέση να κατανοούν το μήνυμα που μεταφέρει το σήμα της φωνής του ομιλητή. Σήμερα καταβάλλονται προσπάθειες να αναπτυχθούν τέτοια συστήματα για να λειτουργούν σε εργασιακά περιβάλλοντα όπου χρησιμοποιούνται υπογλώσσες σε συγκεκριμένα θεματικά πεδία, όπου η χρήση, δηλαδή, του γλωσσικού συστήματος είναι περιορισμένη τόσο σε λεξιλογικό όσο και μορφοσυντακτικό επίπεδο.

Σε σχέση με την Ελληνική γλώσσα, θα πρέπει να επισημανθεί ότι έχουν ήδη αναπτυχθεί εργαλεία αναγνώρισης φωνής, σύνθεσης φωνής και αυτόματης φωνητικής μεταγραφής, τα οποία αξιοποιούνται στο πλαίσιο εκπαιδευτικών εφαρμογών για την υποστήριξη της διδασκαλίας της Ελληνικής. Παράλληλα με τη συνεχή προσπάθεια βελτίωσης των ήδη υπάρχοντων συστημάτων,



έχουν ξεκινήσει προσπάθειες για την ανάπτυξη εργαλείων ελέγχου του τόνου και του επιτονισμού, οι οποίες προς το παρόν κινούνται σε ερευνητικό επίπεδο.

Η αυτόματη αξιολόγηση της προφοράς με τη χρήση εργαλείων τεχνολογίας φωνής ενταγμένων στο πλαίσιο εκπαιδευτικών λογισμικών παρέχει τη δυνατότητα εξατομίκευσης των γλωσσικών αναγκών του χρήστη, την παροχή της κατάλληλης ανατροφοδότησης, την ευκολία να δουλεύει με τους δικούς του ρυθμούς. Τα παραπάνω πλεονεκτήματα, σε συνδυασμό με τις συνεχείς βελτιώσεις αυτών των εργαλείων επιτρέπουν ίσως τον ισχυρισμό ότι στο μέλλον η εκπαιδευτική αξιοποίηση τεχνολογιών φωνής (με την απαραίτητη βεβαίως συνθήκη ότι αξιοποιούνται με τον κατάλληλο τρόπο και από το κατάλληλο ανθρώπινο δυναμικό) θα συμβάλει στην υποστήριξη της γλωσσικής διδασκαλίας και στην κατάκτηση του φωνητικού-φωνολογικού συστήματος της ξένης γλώσσας.

## ΒΙΒΛΙΟΓΡΑΦΙΑ

- Bernstein, L. & Christian, B. (1996), For speech perception by humans or machines, three senses are better than one, *Proceedings of the International Conference on Spoken Language Processing*, October, 1477-1480
- Bernstein, J. & Horacio, F. (1996), Speech recognition by computer, In N. Lass (Ed.), *Principles of experimental phonetics*, St. Louis: Mosby, 408-434
- Celce-Murcia, M., Brinton, D.M. & Goodwin, J.M. (1996), *Teaching Pronunciation*. Cambridge: Cambridge University Press
- Chapelle, C. (1997), CALL in the year 2000: Still in search of research paradigms? *Language Learning & Technology*, 1(1), 19-43
- Crystal, D. (1981), *Clinical linguistics*, New York: Harper Press.
- Holland, M. (1995), The case for intelligent CALL. In M. Holland, J. D. Kaplan, & M. R. Sams (Eds.), *Intelligent language tutors: Theory shaping technology*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Καραγιάννης, Γ. (2001), Δυνατότητες συστημάτων πολυμέσων για εκπαιδευτικές εφαρμογές-Τεχνολογικές τάσεις και εργαλεία, *Πρακτικά Ημερίδας «Γλωσσική εκπαίδευση στην Κοινωνία της Πληροφορίας»*, Ινστιτούτο Επεξεργασίας του Λόγου, 24 Οκτωβρίου, 9-19
- Markham, D. & Nagano-Madsen, Y. (1996), *Proceeding of the International Conference on Spoken Language Processing*, Oct., 1473-1476.
- Salaberry, M.R. (1996), A theoretical foundation for the development of pedagogical tasks in computer mediated communication, *CALICO journal*, 14 (1), 5-34
- Steinhauer, G., Bakamidis, S., Charalabopoulou, F. & Carayannis, G. (2001), Speech tools in the educational process, *Proceedings of the Euronoise Conference 2001*, 14-17 January, 754-757
- Σταϊνχάουερ, Γ., Μπακαμίδης Σ., Καραγιάννης Γ., Αντωνίου Ι., Χαραλαμποπούλου, Φ., Οικονόμου Ν. (2001), Νέες τεχνολογίες και υπολογιστικά εργαλεία στη γλωσσική εκπαίδευση, *Πρακτικά Ημερίδας «Γλωσσική εκπαίδευση στην Κοινωνία της Πληροφορίας»*, Ινστιτούτο Επεξεργασίας του Λόγου, 24 Οκτωβρίου, 55-63
- Warschauer M. (1996), Computer-assisted language learning: An introduction, In S. Fotos (Ed.), *Multimedia language teaching*, 3-20, Logos International: Tokyo