

Ψηφιακά εκπαιδευτικά παιχνίδια τεχνολογιών επεξεργασίας φυσικής γλώσσας στη διάθεση του εκπαιδευτικού: το παράδειγμα του «εκατομμυριούχου» στο μάθημα της γλώσσας

Μόνικα Γαβριηλίδου
ITYE «Διόφαντος»
Monica.Gavrielidou@cti.gr

Μαρία Φουντάνα
ITYE «Διόφαντος»
fountana@cti.gr

Αριστείδης Βαγγελάτος
ITYE «Διόφαντος»
vagelat@cti.gr

Περίληψη

Η παρούσα εργασία αφορά στην παρουσίαση ενός από τα εκπαιδευτικά ψηφιακά παιχνίδια που υλοποιούνται στο πλαίσιο του έργου σχεδιασμού και δημιουργίας γλωσσικών παιχνιδιών με τεχνολογίες επεξεργασίας φυσικής γλώσσας (natural language processing) «Λεξιπαιγνιο», ενός έργου που φιλοδοξεί να διερευνήσει τις δυνατότητες των τεχνολογιών αυτών στον σχεδιασμό εκπαιδευτικών παιχνιδιών.

Πιο συγκεκριμένα, το παιχνίδι «Ποιος θέλει να γίνει εκατομμυριούχος» αφορά εκπαιδευτική παραλλαγή του γνωστού παιχνιδιού, όπου με γρήγορους ρυθμούς και παιγνιώδη τρόπο ο μαθητής εξασκείται σε θέματα δομής και χρήσης της νέας ελληνικής (ορθογραφία, γραμματική, γενικό λεξιλόγιο κ.α.), καθώς και σε γνώσεις και ορολογία σε διάφορα μαθησιακά αντικείμενα (π.χ. στο μάθημα της Γεωγραφίας). Η καινοτομία των παιχνιδιών που αναπτύσσονται στο πλαίσιο του «Λεξιπαιγνίου» συνίσταται στον δυναμικό, αυτοματοποιημένο τρόπο παραγωγής ερωτήσεων αλλά και μη-ορθών εναλλακτικών απαντήσεων, όπου υπάρχει η δυνατότητα πολλαπλής επιλογής, με διαβαθμισμένη δυσκολία, για διαφορετικά γνωστικά αντικείμενα, αξιοποιώντας σώμα κειμένων που περιλαμβάνει τα σχολικά εγχειρίδια και άλλες οντότητες.

Λαμβάνοντας υπόψη ότι η διδασκαλία της Γλώσσας στο σχολείο συχνά περιορίζεται σε τυποποιημένους, μηχανιστικούς τρόπους εκμάθησης κανόνων και λεξιλογίου που μπορεί και να μην εντάσσονται ενίοτε σε κειμενικά περιβάλλοντα, ο «εκατομμυριούχος» φιλοδοξεί να αποτελέσει εργαλείο – βοηθό στα χέρια του εκπαιδευτικού αλλά και μια παιγνιώδη εναλλακτική στη διάθεση των μαθητών, με στόχο τη βελτίωση του γλωσσικού και γενικότερα του μαθησιακού επιπέδου τους με ευχάριστο τρόπο.

Στην παρουσίαση του παιχνιδιού περιλαμβάνονται η περιγραφή και η στοχοθεσία του, ο τρόπος κάλυψης διαφορετικών επιπέδων γνώσεων των μαθητών, οι δυνατότητες του εκπαιδευτικού να προσαρμόσει το περιεχόμενο του παιχνιδιού ανάλογα με τη σχολική τάξη και τη διδακτική ενότητα ή το γραμματικό φαινόμενο που επιθυμεί να καλύψει, στοιχεία μεθοδολογίας και αρχιτεκτονικής του ψηφιακού παιχνιδιού, παραδείγματα εφαρμογής κλπ.

Λέξεις-κλειδιά: ΤΠΕ στην εκπαίδευση, παιγνιώδης μάθηση, επεξεργασία φυσικής γλώσσας, ψηφιακά εκπαιδευτικά παιχνίδια, mini games, Γλώσσα

1. Εισαγωγή – Αναγκαιότητα

Η διδασκαλία του μαθήματος της γλώσσας στους μαθητές της πρωτοβάθμιας και δευτεροβάθμιας εκπαίδευσης προωθεί την κειμενοκεντρική προσέγγιση στη μελέτη των

γραμματικών φαινομένων (Ιορδανίδου, 2007). Στο πλαίσιο αυτό τα γραμματικά φαινόμενα μελετώνται ενταγμένα σε κειμενικά περιβάλλοντα και οι μαθητές εξασκούνται στην αναγνώριση της λειτουργίας των γραμματικών δομών των μηχανισμών της γλώσσας. Αντίστοιχα, στο πλαίσιο της διδασκαλίας του λεξιλογίου οι λέξεις δεν αντιμετωπίζονται μεμονωμένα, αλλά εντάσσονται στο κείμενο και μελετώνται ως δομικά στοιχεία του λόγου. Ταυτόχρονα, τα σχετικά προγράμματα σπουδών επισημαίνουν την ανάγκη για εξεύρεση παιγνιωδών λύσεων στο μάθημα της γλώσσας. Οι παιγνιώδεις αυτές λύσεις θα πρέπει αφενός να είναι παιδαγωγικά άρτιες, εστιασμένες στους διδακτικούς στόχους του μαθήματος, αφετέρου να είναι ελκυστικές και πρωτότυπες έτσι ώστε να διατηρούν αμείωτο το ενδιαφέρον και την κινητοποίηση των μαθητών. Τέτοιες δραστηριότητες θα μπορούν να αξιοποιηθούν για την εξάσκηση των μαθητών σε αυτά που έχουν ήδη διδαχθεί.

Στο πεδίο της τεχνολογίας η τεχνητή νοημοσύνη (Artificial Intelligence) τις τελευταίες δεκαετίες σημειώνει μεγάλη ανάπτυξη και βρίσκει εφαρμογές σε πολλούς τομείς της καθημερινότητάς μας. Αν και η αξιοποίηση της τεχνητής νοημοσύνης στα ψηφιακά παιχνίδια ψυχαγωγίας είναι ευρέως διαδεδομένη, στα εκπαιδευτικά ψηφιακά παιχνίδια παραμένει περιορισμένη (Picca, Jaccard & Eberle 2015). Οι τεχνολογίες επεξεργασίας φυσικής γλώσσας - ΕΦΓ (Natural Language Processing) αποτελούν σημαντικό κλάδο της τεχνητής νοημοσύνης που εστιάζουν στην αλληλεπίδραση υπολογιστή και ανθρώπου με μέσο επικοινωνίας την ανθρώπινη γλώσσα. Η συμβολή των τεχνολογιών επεξεργασίας φυσικής γλώσσας μπορεί να αποδειχθεί ιδιαίτερα χρήσιμη στην ανάπτυξη εκπαιδευτικών παιχνιδιών γλώσσας όπου με δυναμικό τρόπο θα παράγονται γλωσσικές ασκήσεις που θα αξιοποιούν συγκεκριμένο σώμα κειμένων, αλλά και άλλους πόρους υποδομών ΕΦΓ.

Το έργο «Λεξιπαίγνιο» επιχειρεί να συνδράμει στην προσπάθεια αξιοποίησης των τεχνολογιών επεξεργασίας φυσικής γλώσσας στο μάθημα της ελληνικής γλώσσας αναπτύσσοντας μια σειρά από εκπαιδευτικά mini games. Ένα από αυτά είναι το παιχνίδι «εκατομμυριούχος».

2. Στοχοθεσία και Περιγραφή του παιχνιδιού

Στόχος και ταυτόχρονα καινοτομία των απλών και σύντομων παιχνιδιών (mini games) που αναπτύσσονται στο πλαίσιο του «Λεξιπαίγνιου» αποτελεί ο δυναμικός, αυτοματοποιημένος τρόπος παραγωγής ερωτήσεων αλλά και μη-ορθών εναλλακτικών απαντήσεων, όπου υπάρχει η δυνατότητα πολλαπλής επιλογής. Ταυτόχρονα, ο σχεδιασμός των παιχνιδιών αυτών επιτρέπει την παραγωγή ερωτήσεων διαβαθμισμένης δυσκολίας, για διαφορετικά γνωστικά αντικείμενα, αξιοποιώντας σώμα κειμένων που περιλαμβάνει τα σχολικά εγχειρίδια αλλά και άλλες πηγές (π.χ. Βικιπαίδεια, κτ). Το παιχνίδι γνώσεων «Ποιος θέλει να γίνει εκατομμυριούχος» είναι δημοφιλές τηλεοπτικό παιχνίδι, το οποίο στην ψηφιακή του μορφή δίνει τη δυνατότητα στον εκπαιδευτικό να οργανώσει μέρος της διδακτέας ύλης με τη μορφή ερωτήσεων πολλαπλής επιλογής, με σταδιακή διαβάθμιση της δυσκολίας τους. Το παιχνίδι αυτό επιλέχθηκε ως ένα από μια σειρά μικρών παιχνιδιών (mini games) που αναπτύσσονται στο πλαίσιο του έργου με σκοπό την απόδειξη της ορθότητας της ιδέας (proof of concept) αξιοποίησης τεχνολογιών επεξεργασίας φυσικής γλώσσας στη διδασκαλία γλωσσικών μαθημάτων. Ως βάση επιλέχθηκε το ελεύθερο και ανοιχτό λογισμικό (game engine) Godot (<https://godotengine.org/>), όπου έγιναν οι παρακάτω σχεδιαστικές παρεμβάσεις:

- Δυναμικός αυτοματοποιημένος τρόπος παραγωγής ερωτήσεων αλλά και μη ορθών εναλλακτικών απαντήσεων.
- Παραγωγή ερωτήσεων διαβαθμισμένης δυσκολίας.

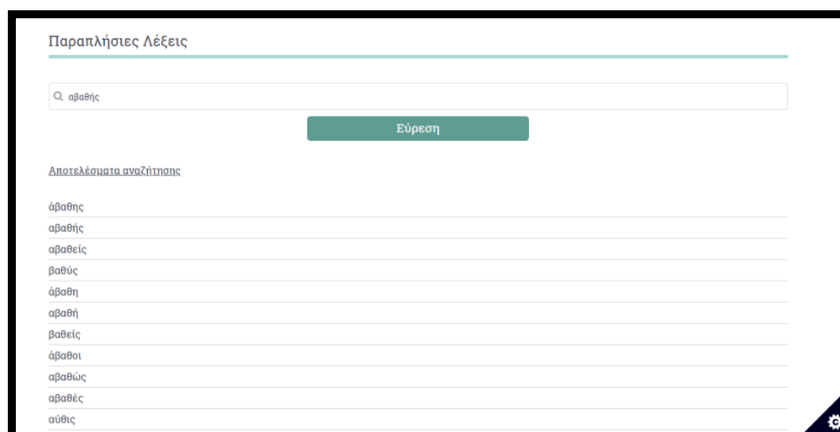
- Αξιοποίηση σώματος κειμένων των σχολικών εγχειριδίων και άλλων κειμένων (Βικιπαίδεια, Φωτόδεντρο κ.α).

Στην παρούσα φάση ανάπτυξης του παιχνιδιού, ως διδακτικός στόχος επιλέχθηκε η εξοικείωση των μαθητών - παικτών με την κλίση των επιθέτων που ανήκουν στην κατηγορία «τριγενή και δικατάληκτα» (Εικόνα 1).



Εικόνα 1. Το περιβάλλον του παιχνιδιού

Οι ερωτήσεις του παιχνιδιού προέκυψαν ύστερα από αυτοματοποιημένη αναζήτηση τριγενών και δικατάληκτων επιθέτων στα βιβλία Γλώσσας και Γραμματικής της Νέας Ελληνικής Γλώσσας για Α', Β', Γ' Γυμνασίου και Ε' & ΣΤ' Δημοτικού καθώς και στα βιβλία της Γεωγραφίας Α' & Β' Γυμνασίου. Η παραγωγή των ερωτήσεων από το παραπάνω υλικό προέκυψε με την απομάκρυνση/απόκρυψη της σωστής λέξης ως «σωστή απάντηση» και την παραγωγή μη ορθών εναλλακτικών απαντήσεων με τα εργαλεία εύρεσης παραπλήσιων λέξεων ή σημασιολογικής ομοιότητας (Εικόνα 2) που έχουν αναπτυχθεί.



Εικόνα 2. Παραγωγή εναλλακτικών απαντήσεων με το εργαλείο εξεύρεσης παραπλήσιων λέξεων

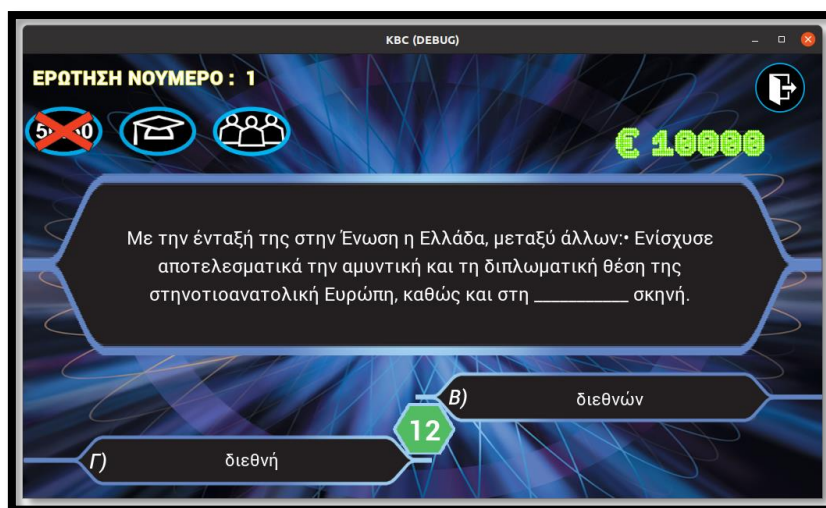
Η «σωστή απάντηση» καθώς και τρεις μη ορθές εναλλακτικές απαντήσεις τροφοδότησαν τη λίστα των πιθανών απαντήσεων σε κάθε ερώτηση του παιχνιδιού (Εικόνα 3).

Ερώτηση	Απάντηση Α	Απάντηση Β	Απάντηση Γ	Απάντηση Δ
1. Πόσο είναι ο μέσος όρος των ετήσιων εισοδημάτων των νοικοκυριών στην Ελλάδα;	α) 10.000€	β) 20.000€	γ) 40.000€	δ) 80.000€
2. Ποιο είναι το μεγαλύτερο ποσοστό των εργαζομένων στην Ελλάδα που εργάζονται στον τομέα των υπηρεσιών;	α) 70%	β) 80%	γ) 90%	δ) 95%
3. Ποιο είναι το μεγαλύτερο ποσοστό των εργαζομένων στην Ελλάδα που εργάζονται στον τομέα της βιομηχανίας;	α) 10%	β) 20%	γ) 30%	δ) 40%
4. Ποιο είναι το μεγαλύτερο ποσοστό των εργαζομένων στην Ελλάδα που εργάζονται στον τομέα της γεωργίας;	α) 1%	β) 2%	γ) 3%	δ) 4%
5. Ποιο είναι το μεγαλύτερο ποσοστό των εργαζομένων στην Ελλάδα που εργάζονται στον τομέα της υγείας;	α) 5%	β) 10%	γ) 15%	δ) 20%
6. Ποιο είναι το μεγαλύτερο ποσοστό των εργαζομένων στην Ελλάδα που εργάζονται στον τομέα της εκπαίδευσης;	α) 3%	β) 5%	γ) 7%	δ) 10%
7. Ποιο είναι το μεγαλύτερο ποσοστό των εργαζομένων στην Ελλάδα που εργάζονται στον τομέα της πληροφορικής;	α) 2%	β) 4%	γ) 6%	δ) 8%
8. Ποιο είναι το μεγαλύτερο ποσοστό των εργαζομένων στην Ελλάδα που εργάζονται στον τομέα της ενέργειας;	α) 1%	β) 2%	γ) 3%	δ) 4%
9. Ποιο είναι το μεγαλύτερο ποσοστό των εργαζομένων στην Ελλάδα που εργάζονται στον τομέα της μεταφοράς;	α) 1%	β) 2%	γ) 3%	δ) 4%
10. Ποιο είναι το μεγαλύτερο ποσοστό των εργαζομένων στην Ελλάδα που εργάζονται στον τομέα της κατασκευής;	α) 1%	β) 2%	γ) 3%	δ) 4%
11. Ποιο είναι το μεγαλύτερο ποσοστό των εργαζομένων στην Ελλάδα που εργάζονται στον τομέα της εμπορίας;	α) 1%	β) 2%	γ) 3%	δ) 4%
12. Ποιο είναι το μεγαλύτερο ποσοστό των εργαζομένων στην Ελλάδα που εργάζονται στον τομέα της επικοινωνίας;	α) 1%	β) 2%	γ) 3%	δ) 4%

Εικόνα 3. Βάση αυτοματοποιημένων ερωτήσεων και μη ορθών εναλλακτικών απαντήσεων

Σύμφωνα με τους κανόνες του ψηφιακού παιχνιδιού του «εκατομμυριούχου» ο παίκτης καλείται να απαντήσει σε έντεκα ερωτήσεις, κάθε μία από τις οποίες αντιστοιχεί σε συγκεκριμένο εικονικό χρηματικό ποσό: 10.000€, 20.000€, 40.000€, 80.000€, 160.000€, 320.000€, 640.000€, 1.250.000€, 2.500.000€, 5.000.000€ και 10.000.000€. Η παράμετρος του χρόνου του παιχνιδιού είναι προσαρμοσμένη στους κανόνες του παιχνιδιού και περιορίζεται στα 35'' δευτερόλεπτα για κάθε ερώτηση.

Ο παίκτης επιλέγει τη σωστή απάντηση για κάθε μία από τις έντεκα ερωτήσεις και κερδίζει το αντίστοιχο εικονικό χρηματικό ποσό. Το παιχνίδι τελειώνει όταν απαντηθούν σωστά όλες οι ερωτήσεις, όταν μια ερώτηση απαντηθεί λάθος ή όταν δεν απαντηθεί μια ερώτηση στο χρονικό περιθώριο των 35 δευτερολέπτων. Το παιχνίδι ακολουθεί τη διαδικασία βαθμολόγησης στην οποία συμμετέχουν τρεις συμπληρωματικές βοήθειες που υπάρχουν ενσωματωμένες στην αρχική έκδοση του ψηφιακού παιχνιδιού. Η βοήθεια του 50-50, εξαφανίζει δύο από τις τέσσερις υποψήφιες απαντήσεις (Εικόνα 4).



Εικόνα 4. Η βοήθεια 50-50

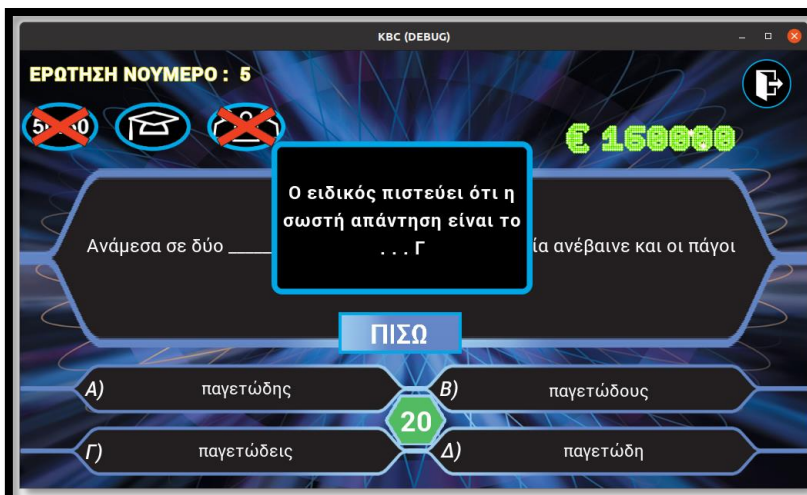
Από τις απαντήσεις που μένουν, μία είναι πάντα η σωστή απάντηση ενώ επιλέγεται τυχαία και μία δεύτερη απάντηση από τις μη ορθές απαντήσεις.

Η βοήθεια του κοινού εμφανίζει ραβδόγραμμα τεσσάρων ράβδων που αντιστοιχούν σε κάθε απάντηση, δίνοντας ένα συγκεκριμένο αριθμό πιθανοτήτων για κάθε απάντηση, σύμφωνα με τη γνώμη του εικονικού κοινού του παιχνιδιού (Εικόνα 5).



Εικόνα 5. Η βοήθεια του κοινού

Τέλος, η βοήθεια της γνώμης ενός ειδικού έχει ως αποτέλεσμα να προτείνεται στον παίκτη μια μόνο απάντηση. Σύμφωνα με τον σχεδιασμό του παιχνιδιού η απάντηση αυτή είναι πάντα και η σωστή (Εικόνα 6).



Εικόνα 6. Η βοήθεια του ειδικού

Το επίπεδο δυσκολίας των ερωτήσεων του παιχνιδιού είναι διαβαθμισμένο και προσδιορίζεται με βάση το σύνολο των κειμένων των σχολικών βιβλίων των τάξεων του Δημοτικού και του Γυμνασίου όπου αρχικά συναντάται το γραμματικό φαινόμενο που εξασκείται στο παιχνίδι. Ταυτόχρονα, με την παραδοχή ότι το ίδιο γραμματικό φαινόμενο συναντάται και σε πιο σύνθετα κείμενα που απευθύνονται σε μεγαλύτερες τάξεις μαθητών, τις αρχικές – απλούστερες- ερωτήσεις ακολουθούν συνθετότερες, με αποτέλεσμα να αλλάζει το επίπεδο δυσκολίας του παιχνιδιού. Έτσι λοιπόν, στο πλαίσιο του παιχνιδιού ισχύει ο ακόλουθος σχεδιαστικός κανόνας: τα πρώτα ερωτήματα βασίζονται σε προτάσεις από κείμενα μικρότερων τάξεων και προοδευτικά επιλέγονται προτάσεις μεγαλύτερων τάξεων.

3. Ο ρόλος του εκπαιδευτικού

Το περιβάλλον του παιχνιδιού δίνει τη δυνατότητα στον εκπαιδευτικό να προσαρμόσει το περιεχόμενο του παιχνιδιού ανάλογα με τη σχολική τάξη, τη διδακτική ενότητα ή το γραμματικό φαινόμενο που επιθυμεί να καλύψει. Για παράδειγμα, για την εξάσκηση στα τριγενή και δικατάληκτα επίθετα σε μαθητές της Ε' δημοτικού ο εκπαιδευτικός θα κάνει τη σχετική επιλογή στο αρχικό μενού του παιχνιδιού. Αντίστοιχα, για την εξάσκηση μεγαλύτερων ηλικιακά μαθητών στο ίδιο φαινόμενο, ο εκπαιδευτικός θα μπορεί να επιλέξει μια εκδοχή του παιχνιδιού που θα παρέχει ερωτήματα μεγαλύτερης δυσκολίας. Εξ άλλου, ο σχεδιασμός του παιχνιδιού είναι τέτοιος ώστε να επιτρέπει στον εκπαιδευτικό να παίρνει πρωτοβουλίες που αφορούν την εμβάθυνση και σε άλλα γραμματικά φαινόμενα ή ακόμη και τον εμπλουτισμό του σώματος κειμένων που θα επεξεργάζεται το περιβάλλον του παιχνιδιού. Κατά συνέπεια, εφόσον το επιθυμεί, ο εκπαιδευτικός θα μπορεί να δημιουργήσει νέες ερωτήσεις που θα αφορούν στοιχεία του λεξιλογίου ή της γραμματικής, ή ακόμη και να προσθέσει στο σώμα κειμένων του παιχνιδιού κείμενα της επιλογής του.

4. Μεθοδολογία και Αρχιτεκτονική των υποδομών του παιχνιδιού

Για την υλοποίηση του παιχνιδιού, στο πλαίσιο του έργου «Λεξιπαίγνιο», υλοποιήθηκαν/επεκτάθηκαν μια σειρά από υποδομές και πόρους επεξεργασίας φυσικής γλώσσας (Tsalidis, Fountana, Gavrielidou, 2020). Έτσι, προηγήθηκαν τα παρακάτω:

4.1 Συγκρότηση σώματος εκπαιδευτικών κειμένων

Με βάση τόσο τον «τεχνικό σχεδιασμό» του έργου, όσο και τις πραγματικές ανάγκες του, συγκροτήθηκε σώμα εκπαιδευτικών κειμένων. Πιο συγκεκριμένα, με βάση κατά κύριο λόγο τα σχολικά βιβλία, και δευτερευόντως άλλες πηγές, συγκροτήθηκε σώμα κειμένων:

A) Γεωγραφίας, B) Βιολογίας, Γ) Θρησκευτικών, Δ) Μουσικής και E) Οικιακής Οικονομίας για την Α' Γυμνασίου.

Το υλικό, αναρτήθηκε σε κοινόχρηστο χώρο της ομάδας έργου στο «σύννεφο», ώστε να υπάρχει δυνατότητα άμεσης πρόσβασης από όλους.

Ειδικά για το μάθημα της Γεωγραφίας - Γεωλογίας, το οποίο αποτελεί και το δοκιμαστικό πεδίο για το έργο, αποφασίστηκε να εμπλουτιστεί το σώμα κειμένων περαιτέρω έτσι ώστε να αποτελείται από:

A) τα σχολικά βιβλία της Γεωγραφίας – Γεωλογίας της Α' τάξης Γυμνασίου (βιβλίο μαθητή, Τετράδιο Εργασιών και Βιβλίο Εκπαιδευτικού¹),

B) τα σχολικά βιβλία της Γεωγραφίας – Γεωλογίας της Β' τάξης Γυμνασίου (βιβλίο μαθητή, Τετράδιο Εργασιών και Βιβλίο Εκπαιδευτικού),

Γ) τα σχολικά βιβλία της Γεωγραφίας – Γεωλογίας του Δημοτικού (Ε' και ΣΤ' τάξεις), αλλά και Δ) το ΑΠΣ-ΔΕΠΠΣ (αναλυτικό πρόγραμμα σπουδών και διαθεματικό ενιαίο πλαίσιο προγραμμάτων σπουδών) για το μάθημα της Γεωγραφίας – Γεωλογίας (Δημοτικού Γυμνασίου) και τέλος

E) τα εγχειρίδια λογισμικού (ένα για μαθητές και ένα για εκπαιδευτικούς) του Παιδαγωγικού Ινστιτούτου, για λογισμικό Γεωγραφίας που είχε αναπτυχθεί και διανεμηθεί στα σχολεία καθώς και

¹ Το βιβλίο Εκπαιδευτικού ήταν σε μη επεξεργάσιμη μορφή (εικόνα) και χρειάστηκε να περάσει από λογισμικό οπτικής αναγνώρισης χαρακτήρων (OCR) και κάποια επεξεργασία διόρθωσης λαθών, ώστε να γίνει αξιοποιήσιμο.

ΣΤ) υλικό από τον εκπαιδευτικό κόμβο "Ίφιγένεια" με κείμενα που σχετίζονται με το μάθημα αυτό).

Ο στόχος της συγκρότησης του σώματος κειμένων ήταν η αυτόματη εξαγωγή θεματικού λεξιλογίου και άλλης δομημένης κειμενικής πληροφορίας, που θα επιτρέπει τη δημιουργία παιχνιδιών και ερωτήσεων ανά θεματικό πεδίο.

4.2 Περιγραφή συντακτικών σχημάτων για αναγνώριση όρων

Για την αναγνώριση όρων, ήταν απαραίτητο αρχικά να αναλυθεί η δομή των βιβλίων της γεωγραφίας σε μια προσπάθεια να εξευρεθούν «μοτίβα» (patterns) που θα βοηθήσουν στην αυτοματοποίηση της αναγνώρισής τους. Το αποτέλεσμα της ανάλυσης έδωσε αρκετή πληροφορία για την αξιοποίηση του περιεχομένου, π.χ. μπορούν να εξαχθούν «ορισμοί» με βάση ένα συγκεκριμένο μοτίβο, αλλά και προβλήματα π.χ. ακόμα και στο ίδιο βιβλίο το μοτίβο για τους ορισμούς συχνά αλλάζει!

Στη συνέχεια περιγράφηκαν κανόνες (συντακτικών σχημάτων) βάσει των οποίων οργανώνεται η ορολογική πληροφορία των κειμένων αυτών βάσει των οποίων διατυπώνονται στη συνέχεια οι γλωσσικές προδιαγραφές που στηρίζουν την αυτόματη εξαγωγή ορολογίας και την αυτόματη οργάνωσή της σε ταξινομίες όρων. Ένας τέτοιος κανόνας είναι ο ακόλουθος:

Περιγραφή 1^{ου} κανόνα: Ο κανόνας περιγράφει μια ονοματική φράση όπου ένα επίθετο ακολουθείται από ένα ουσιαστικό που συμφωνεί με το προηγούμενο επίθετο σε γένος, αριθμό και πτώση.

Παραδείγματα: γεωγραφικό διαμέρισμα, γεωγραφικό μήκος, μεγαλύτερη πόλη, ανατολική περιοχή

ΚΑΝΟΝΑΣ 1 1:

```
[ΚΑΝΟΝΑΣ="ΠΟΛΥΛΕΚΤΙΚΟ_1", ΜτΛ=ΟΥΣΙΑΣΤΙΚΟ] ==>
\
[ΜτΛ=(ΕΠΙΘΕΤΟ), ΣΥΜΦΩΝΙΑ=(ΓΕΝΟΣ, ΑΡΙΘΜΟΣ, ΠΤΩΣΗ)],
[ΜτΛ=(ΟΥΣΙΑΣΤΙΚΟ), ΣΥΜΦΩΝΙΑ=(ΓΕΝΟΣ, ΑΡΙΘΜΟΣ, ΠΤΩΣΗ)]
/
;
```

Εφαρμογή κανόνα:

γεωγραφικό διαμέρισμα =>

```
\
γεωγραφικό/{ΕΠΙΘΕΤΟ,ΑΡΣΕΝΙΚΟ,ΟΝΟΜΑΣΤΙΚΗ,ΕΝΙΚΟΣ}
διαμέρισμα/{ΟΥΣΙΑΣΤΙΚΟ,ΑΡΣΕΝΙΚΟ,ΟΝΟΜΑΣΤΙΚΗ,ΕΝΙΚΟΣ}
/
```

Οι κανόνες αυτοί «μεταφράστηκαν» στη συνέχεια με βάση τον ειδικό φορμαλισμό «Κανών» (Tsalidis, Pantazara, Minos, 2010; Vagelatos, Gavrielidou, Fountana, 2021; Gakis, Panagiotakopoulos, Sgarbas, 2015). Ο φορμαλισμός «Κανών», εκτός από τη δυνατότητα ορισμού μορφοσυντακτικών σχημάτων ευαίσθητων περιεχομένου (context sensitive constructs) που λαμβάνουν υπόψη το συμφραστικό τους περιβάλλον, έχει τη δυνατότητα να αξιοποιεί επιμέρους χαρακτηριστικά των λέξεων (μορφοσυντακτικά, υφολογικά κτλ.).

4.3 Γραμματικά Φαινόμενα και Γλωσσικά Λάθη

Παράλληλα με τις υποδομές ΕΦΓ, πραγματοποιήθηκε καταγραφή και κατηγοριοποίηση των γραμματικών φαινομένων στο μάθημα της Γλώσσας για την υποχρεωτική εκπαίδευση με βάση τα Προγράμματα Σπουδών (Διαθεματικό Ενιαίο Πλαίσιο Προγραμμάτων Σπουδών, Δ.Ε.Π.Π.Σ.) και τα διδακτικά εγχειρίδια του μαθήματος της Γλώσσας (Δημοτικό) και Νεοελληνικής Γλώσσας (Γυμνάσιο). Κρίθηκε σκόπιμο η συγκεκριμένη ανάλυση να πραγματοποιηθεί για το διδακτικό περιεχόμενο του Δημοτικού και του Γυμνασίου (όχι για το Λύκειο) τόσο διότι ολόκληρη η ύλη της Γραμματικής εξαντλείται στο Δημοτικό, όσο και λόγω των ιδιαιτεροτήτων του προγράμματος της σχολικής βαθμίδας του Λυκείου.

Αντίστοιχα, ορίσθηκε το «γλωσσικό λάθος» αλλά και το αντίστοιχο γλωσσικό/γραμματικό φαινόμενο που πρέπει κανείς να εστιάσει για την κατάλληλη μεθοδολογία μάθησης που θα εφαρμοστεί.

4.4 Αρχιτεκτονική του συστήματος

Αναλυτικότερα, τα βήματα που ακολουθήθηκαν για την ανάλυση και εξαγωγή των γλωσσικών πληροφοριών από τα κείμενα είναι τα ακόλουθα:

1. **Συγκρότηση σώματος κειμένων.** Το αρχικό σώμα κειμένων αποτελείται κυρίως από τα σχολικά βιβλία που αναφέρθηκαν παραπάνω.

2. **Κατάτμηση των κειμένων σε περιόδους, αναγνώριση σημείων στίξης και μικρότερων ενοτήτων στο πλαίσιο των περιόδων (tokenisation).** Η διαδικασία αυτή επιτρέπει την αναγνώριση των ενοτήτων μέσα στις οποίες εντοπίζονται τα γεγονότα σχολιασμού (π.χ. περίοδος, πρόταση).

3. **Αυτόματη μορφοσυντακτική αναγνώριση των λέξεων του κειμένου και λημματοποίηση.**

- Η **λημματοποίηση** αφορούσε την απόδοση σε κάθε λέξη του λημματικού τύπου που την αντιπροσωπεύει (π.χ. «Βρείτε στην υδρόγειο σφαίρα τον Ισημερινό» → Λημματοποίηση: ΒΡΙΣΚΩ ΣΤΗΝ ΥΔΡΟΓΕΙΟΣ ΣΦΑΙΡΑ Ο ΙΣΗΜΕΡΙΝΟΣ).

- Κατά τη **μορφοσυντακτική** αναγνώριση αποδόθηκαν σε κάθε λέξη επισημειώσεις (tags) που αφορούν το μέρος του λόγου (π.χ. ουσιαστικό, επίθετο, ρήμα).

4. **Εξαγωγή μονολεκτικών και πολυλεκτικών όρων:** α) μέσω κανόνων που αφορούν την αναγνώριση ονοματικών φράσεων (ΟΦ) και ρηματικών φράσεων (ΡΦ) και β) με στατιστική ανάλυση που υπολογίζει λέξεις που απαντούν συχνά στα ίδια συμφραζόμενα. Για την στατιστική ανάλυση χρησιμοποιήθηκαν δύο ελεύθερες εφαρμογές ανοικτού κώδικα: το **KNIME** (<https://www.knime.com/>), μια πλατφόρμα για ανάλυση κειμένων, και το **KH-Coder** (<https://kncoder.net/en/>), μια εφαρμογή για ποσοτική (quantitative) ανάλυση περιεχομένου και εξαγωγή πληροφορίας (text mining).

5. **Δημιουργία λεξικών.** Πρόκειται για δύο (2) λεξικά:

- **Λεξικό πολυλεκτικών όρων**, που περιλαμβάνει πολυλεκτικά λήμματα και φράσεις, το οποίο αποτελεί το αποτέλεσμα της τελικής επιμέλειας των όρων που προέκυψαν από το βήμα 4, από τους γλωσσολόγους/φιλόλογους.

- **Λεξικό μονολεκτικών όρων**, που περιλαμβάνει μονολεκτικά λήμματα, το οποίο αποτελεί το αποτέλεσμα της τελικής επιμέλειας των όρων που προέκυψαν από το βήμα 4, από τους γλωσσολόγους/φιλόλογους.

6. **Δημιουργία ταξινομίας.** Η ταξινομία της Γεωγραφίας.

7. **Δημιουργία και εφαρμογή κανόνων.** Για τη δημιουργία των κανόνων χρησιμοποιήθηκε ο φορμαλισμός «ΚΑΝΩΝ».

8. **Δημιουργία Παιχνιδιού.** Με βάση τις υποδομές, έγινε αυτόματη επιλογή προτάσεων από το βιβλίο της Γεωγραφίας. Πάνω σε αυτές αναγνωρίζονται ανάλογα με τις ανάγκες γραμματικά φαινόμενα (π.χ. τριγενή και δικατάληκτα, σύνθετες λέξεις) και προσφέρονται στον χρήστη για επίλυση με τη μορφή και τις απαιτήσεις του παιχνιδιού.

5. Συμπεράσματα

Τα μέχρι στιγμής αποτελέσματα του έργου, μεταξύ των οποίων είναι και το παιχνίδι «ποιος θέλει να γίνει εκατομμυριούχος» που παρουσιάζεται στην παρούσα εργασία, δείχνουν ιδιαίτερα ενθαρρυντικά. Ιδιαίτερα η δυναμική παραγωγή των παιχνιδιών, αλλά και η δυνατότητα που θα έχει ο εκπαιδευτικός να προσαρμόσει τα παιχνίδια αυτά στις ανάγκες του, θα του δίνει την ευκαιρία να εμπλουτίσει το μάθημα με παιγνιώδεις εργασίες που θα προσελκύσουν περισσότερο τον μαθητή, από τις σημερινές λίγες και στατικές ασκήσεις που περιλαμβάνουν τα σχολικά βιβλία ή το υλικό που έχει συγκεντρώσει ο εκπαιδευτικός στη διάρκεια του χρόνου. Στόχο της ομάδας έργου, ως επόμενο βήμα, αποτελεί η αξιολόγηση των αποτελεσμάτων του έργου σε πραγματικές συνθήκες, μέσα και έξω από την τάξη.

Το έργο «Λεξιπαίγνιο» υλοποιείται στο πλαίσιο της Δράσης ΕΡΕΥΝΩ – ΔΗΜΙΟΥΡΓΩ - ΚΑΙΝΟΤΟΜΩ και συγχρηματοδοτήθηκε από το Ευρωπαϊκό Ταμείο Περιφερειακής Ανάπτυξης (ΕΤΠΑ) της Ευρωπαϊκής Ένωσης και εθνικούς πόρους μέσω του Ε.Π. Ανταγωνιστικότητα, Επιχειρηματικότητα & Καινοτομία (ΕΠΑνΕΚ) (κωδικός έργου: Τ1ΕΔΚ-1ΕΔΚ-05194). Περισσότερες πληροφορίες στην ιστοσελίδα: <http://lexipaignio.cti.gr/>

Βιβλιογραφικές Αναφορές

- Gakis P., Panagiotakopoulos C., Sgarbas K., Tsalidis C., Verykios V. (2015) Construction of a Modern Greek Grammar Checker Through Mnemosyne Formalism. In Ronzhin A., Potapova R., Fakotakis N. (Eds.), *Speech and Computer. SPECOM 2015. Lecture Notes in Computer Science*, vol 9319. Springer.
- Picca, D., Jaccard, D., Eberle, G. (2015). Natural Language Processing in Serious Games. *International Journal of Serious Games*, 2 (3) pp. 77-97. Retrieved from [988f1b221c1b8a1fd8c098c3addc6e930f2d.pdf \(semanticscholar.org\)](http://www.semanticscholar.org/paper/988f1b221c1b8a1fd8c098c3addc6e930f2d)
- Tsalidis, C., Fountana, M., Gavrielidou, M., Vagelatos, A., Stamatopoulos, J. (2020). Implementing language games with NLP tools: The Greek case. In Spyropoulos, C. & Varlamis I. (Eds.), *11th Hellenic Conference on Artificial Intelligence (SETN2020), Games and AI*. Athens, Greece. Retrieved from [11th Hellenic Conference on Artificial Intelligence | ACM Other conferences](http://www.setn2020.org/)
- Tsalidis, C., Pantazara, M., Minos, P. & Mantzari, E. (2010). NLP tools for lexicographic applications. In *Modern Greek, eLexicography in the 21st century: new challenges, new applications, Proceedings of elex 2009*, Louvain-la-Neuve 22-24 October 2009, Cahiers du CENTAL No 7, Université Catholique de Louvain, pp. 457-462.
- Vagelatos, A., Gavrielidou, M., Fountana, M., Tsalidis, C. (2021). Utilizing NLP tools for the creation of school educational games. In Auer, M. & Ruutmann, T. (Eds.) *Educating Engineers for Future Industrial Revolutions*. «Advances in Intelligent Systems and Computing» book series, Springer Retrieved from [Educating Engineers for Future Industrial Revolutions - Proceedings of the 23rd International Conference on Interactive Collaborative Learning \(ICL2020\), Volume 1 | Michael E. Auer | Springer](http://www.springer.com/9783030644444)

Ιορδανίδου, Α. (2007). Κειμενοκεντρικές Προσεγγίσεις του Σχολικού Εγγραμματος: Κείμενο, Συμφραζόμενα, Γραμματική. Στο Η. Ματσαγγούρας (Επιμ.), *Σχολικός Εγγραμματοςμός. Λειτουργικός, Κριτικός, Επιστημονικός*. Αθήνα: Γρηγόρη.