



Σχολή Ανθρωπιστικών Επιστημών
Ψηφιακές Ανθρωπιστικές Επιστήμες

Μεταπτυχιακή Διπλωματική Εργασία

**Εκπαιδευτικός Σχεδιασμός και Παιδαγωγική
Αξιολόγηση Συστήματος Τεχνητής Νοημοσύνης για
Εξατομικευμένη Μάθηση στην Ισπανόφωνη
Λογοτεχνία**

Θεοφάνης Γεροδήμος

Ιωάννινα, Ιούνιος 2026

Η παρούσα εργασία αποτελεί πνευματική ιδιοκτησία του φοιτητή («συγγραφέας/δημιουργός») που την εκπόνησε. Στο πλαίσιο της πολιτικής ανοικτής πρόσβασης ο συγγραφέας/δημιουργός εκχωρεί στο ΕΑΠ, μη αποκλειστική άδεια χρήσης του δικαιώματος αναπαραγωγής, προσαρμογής, δημόσιου δανεισμού, παρουσίασης στο κοινό και ψηφιακής διάχυσής τους διεθνώς, σε ηλεκτρονική μορφή και σε οποιοδήποτε μέσο, για διδακτικούς και ερευνητικούς σκοπούς, άνευ ανταλλάγματος και για όλο το χρόνο διάρκειας των δικαιωμάτων πνευματικής ιδιοκτησίας. Η ανοικτή πρόσβαση στο πλήρες κείμενο για μελέτη και ανάγνωση δεν σημαίνει καθ' οιονδήποτε τρόπο παραχώρηση δικαιωμάτων διανοητικής ιδιοκτησίας του συγγραφέα/δημιουργού ούτε επιτρέπει την αναπαραγωγή, αναδημοσίευση, αντιγραφή, αποθήκευση, πώληση, εμπορική χρήση, μετάδοση, διανομή, έκδοση, εκτέλεση, «μεταφόρτωση» (downloading), «ανάρτηση» (uploading), μετάφραση, τροποποίηση με οποιονδήποτε τρόπο, τμηματικά ή περιληπτικά της εργασίας, χωρίς τη ρητή προηγούμενη έγγραφη συναίνεση του συγγραφέα/δημιουργού. Ο συγγραφέας/δημιουργός διατηρείτο σύνολο των ηθικών και περιουσιακών του δικαιωμάτων.



Σχολή Ανθρωπιστικών Επιστημών
Ψηφιακές Ανθρωπιστικές Επιστήμες

Μεταπτυχιακή Διπλωματική Εργασία

**Εκπαιδευτικός Σχεδιασμός και Παιδαγωγική
Αξιολόγηση Συστήματος Τεχνητής Νοημοσύνης για
Εξατομικευμένη Μάθηση στην Ισπανόφωνη
Λογοτεχνία**

Θεοφάνης Γεροδήμος

Επιτροπή Επίβλεψης Διπλωματικής Εργασίας

Επιβλέπουσα Καθηγήτρια:
Ελένη Κουγιουμτζίδου
Ε.Α.Π.

Συν-επιβλέπων Καθηγητής:
Παναγιώτης Αρβανίτης
Α.Π.Θ.

Ιωάννινα, Ιούνιος 2026

Ευχαριστίες

*Ευχαριστώ από τα βάθη της καρδιάς μου την καθηγήτριά μου Ελένη Κουγιουμτζίδου
για την καθοδήγησή της και την υπομονή της όλο αυτό το διάστημα.*

*Θερμές ευχαριστίες οφείλω και στη Διευθύντρια του Μεταπτυχιακού Προγράμματος
Σπουδών, καθηγήτρια Ξανθίππη Δημητρούλια, για την εμπιστοσύνη της και την
έμπρακτη στήριξη που έδειξε στο πρόσωπό μου.*

*Στην οικογένειά μου
που βρίσκεται πάντα δίπλα μου να με στηρίζει.*

Περίληψη

Η παρούσα διπλωματική εργασία πραγματεύεται τον εκπαιδευτικό σχεδιασμό και την παιδαγωγική αξιολόγηση ενός προσαρμοστικού συστήματος Τεχνητής Νοημοσύνης για τη διδασκαλία της ισπανικής γλώσσας μέσα από λογοτεχνικά κείμενα. Το σύστημα αξιοποιεί Μεγάλα Γλωσσικά Μοντέλα για την αυτόματη παραγωγή εκπαιδευτικών δραστηριοτήτων προσαρμοσμένων στα επίπεδα του Κοινού Ευρωπαϊκού Πλαισίου Αναφοράς (CEFR), με θεωρητικό υπόβαθρο τη σύγκλιση τριών αξόνων: του κονστρουκτιβισμού (κοινωνικοπολιτισμική θεωρία του Vygotsky, σκαλωσιά, μη-δομημένα προβλήματα), της προσαρμοστικής μάθησης (μοντέλο ARCS) και της Τεχνητής Νοημοσύνης στην εκπαίδευση. Αρχιτεκτονικά, η εφαρμογή ακολουθεί το μοντέλο πελάτη-εξυπηρετητή και ενσωματώνει μηχανική προτροπών που κωδικοποιεί τις αρχές του Σωκρατικού διαλόγου, της σκαλωσιάς και της σταδιακής απόσυρσης, καθώς και υποδομή ασύγχρονης ερευνητικής καταγραφής.

Η αξιολόγηση ακολούθησε μεικτή μεθοδολογία (συγκλίνων παράλληλος σχεδιασμός) με τριγωνισμό δεδομένων, αξιοποιώντας ερωτηματολόγια μαθητών και εκπαιδευτικών, ανοιχτές ερωτήσεις και αρχεία καταγραφής των αλληλεπιδράσεων. Στην πιλοτική μελέτη συμμετείχαν μαθητές και εκπαιδευτικοί, με σώμα κειμένων αποσπάσματα των Neruda, García Márquez και Cortázar, αντίστοιχα για τα επίπεδα A1-A2, B1-B2 και C1-C2.

Τα ευρήματα δείχνουν ότι η παιδαγωγική αξία του συστήματος δεν αποτελεί εγγενή ιδιότητα του γλωσσικού μοντέλου, αλλά εξαρτάται από τον σχεδιασμό των προτροπών, την επιλογή του μοντέλου και τη σταθερότητα της υλοποίησης. Η μαθησιακή εμπλοκή ήταν σταθερά υψηλή σε όλα τα επίπεδα, ενώ η συνολική ικανοποίηση παρέμεινε πιο συγκρατημένη, κυρίως λόγω τεχνικών αστοχιών. Αναδείχθηκαν επίσης διαφορές μεταξύ των μοντέλων ως προς τη συνέπεια του Σωκρατικού διαλόγου, η αδυναμία υλοποίησης της σταδιακής απόσυρσης ελλείψει μόνιμης μνήμης, και η σημασία της επιλογής γλώσσας ως παιδαγωγικής παραμέτρου. Συμπερασματικά, το σύστημα αποτιμάται ως αξιόλογο συμπληρωματικό εκπαιδευτικό εργαλείο και όχι ως υποκατάστατο του εκπαιδευτικού.

Λέξεις-κλειδιά: Τεχνητή Νοημοσύνη, Εκπαίδευση, Μεγάλα Γλωσσικά Μοντέλα, Προσαρμοστική Μάθηση, Κονστρουκτιβισμός, Σκαλωσιά, Σωκρατικός Διάλογος, Μηχανική Προτροπών, Ισπανική Λογοτεχνία, CEFR

Abstract

This thesis addresses the instructional design and pedagogical evaluation of an adaptive Artificial Intelligence system for teaching the Spanish language through literary texts. The system leverages Large Language Models to automatically generate educational activities adapted to the levels of the Common European Framework of Reference for Languages (CEFR). Its theoretical foundation lies at the convergence of three axes: constructivism (Vygotsky's sociocultural theory, scaffolding, ill-structured problems), adaptive learning (the ARCS model), and Artificial Intelligence in education. Architecturally, the application follows the client-server model and incorporates prompt engineering that encodes the principles of Socratic dialogue, scaffolding, and gradual fading, as well as an infrastructure for asynchronous research logging.

The evaluation followed a mixed-methods approach (convergent parallel design) with data triangulation, drawing on questionnaires for students and teachers, open-ended questions, and logs of the interactions. The pilot study involved students and teachers, using as its corpus excerpts by Neruda, García Márquez, and Cortázar, corresponding to levels A1-A2, B1-B2, and C1-C2, respectively.

The findings indicate that the pedagogical value of the system is not an inherent property of the language model, but depends on prompt design, model selection, and the stability of the implementation. Learner engagement was consistently high across all levels, whereas overall satisfaction remained more moderate, mainly due to technical shortcomings. The study also revealed differences between the models in the consistency of Socratic dialogue, the inability to implement gradual withdrawal of support in the absence of persistent memory, and the importance of language choice as a pedagogical parameter. In conclusion, the system is assessed as a valuable complementary educational tool rather than a substitute for the teacher.

Keywords: Artificial Intelligence, Education, Large Language Models, Adaptive Learning, Constructivism, Scaffolding, Socratic Dialogue, Prompt Engineering, Spanish Literature, CEFR

Περιεχόμενα

Κεφάλαιο 1: Εισαγωγή.....	1
1.1. Η Ψηφιακή Μεταμόρφωση της Εκπαίδευσης και η Άνοδος της Τεχνητής Νοημοσύνης	1
1.2. Σκοπός της εργασίας	2
1.3. Διάρθρωση της εργασίας.....	2
Κεφάλαιο 2: Θεωρητικό Πλαίσιο και Σύνθεση	4
2.1.Οριοθέτηση της Ανασκόπησης.....	4
2.2. Οι Θεμελιώδεις Αρχές του Κονστрукτιβισμού.....	4
2.2.1. Κριτική Σκέψη και Δημιουργικότητα (Ανάπτυξη Δεξιοτήτων Ανώτερου Επιπέδου).....	5
2.2.2. Η Κοινωνικοπολιτισμική Θεωρία του Vygotsky και τα Γνωστικά Εργαλεία	6
2.2.3. Ενεργός Δόμηση της Γνώσης και τα Μη-Δομημένα Προβλήματα	8
2.3. Προσαρμοστική Μάθηση.....	10
2.3.1. Ορισμός και Μηχανισμοί της Προσαρμοστικής Μάθησης (Adaptive Learning)	10
2.3.2. Μαθησιακά Στυλ και Σχεδιασμός Κινήτρων (Μοντέλο ARCS).....	11
2.3.3. Η Μαθησιακή Εμπλοκή και τα Κίνητρα	12
2.4. Η ΤΝ στην Εκπαίδευση (Edu-AI).....	13
2.4.1. Το Μοντέλο Τριών Επιπέδων (Micro, Meso, Macro).....	13
2.4.2. Η Ενσωμάτωση των Μεγάλων Γλωσσικών Μοντέλων (LLMs)	14
2.4.3. Ηθικές Διαστάσεις: Αλγοριθμική Μεροληψία και Πολιτισμική Ευαισθησία	16
Κεφάλαιο 3: Αρχιτεκτονική Συστήματος	18
3.1. Σχεδιαστικές Αρχές	18
3.1.1. Παιδαγωγική Προσέγγιση της Αρχιτεκτονικής	18
3.1.2. Επιλογή Τεχνολογικού Stack	18
3.2. Αρχιτεκτονική Υψηλού Επιπέδου	20
3.2.1. Το μοντέλο Πελάτη-Εξυπηρετητή	20
3.2.2. Ροή Δεδομένων	21
3.3. Αρχιτεκτονική του Πελάτη (Frontend Modules)	22
3.3.1. Διαχωρισμός των Αρμοδιοτήτων και Κεντρικός Έλεγχος	22

3.3.2. Πολυτροπικότητα του Συστήματος.....	24
3.3.3. Παραγωγή Εκπαιδευτικού Υλικού	24
3.4. Μηχανική Προτροπών (Prompt Enginnering).....	24
3.4.1. Καθοδήγηση έναντι Εκπαίδευσης (Prompt vs Fine Tuning).....	24
3.4.2. Δομή του Συστήματος Προτροπών	25
3.4.3. Επιλογή Γλώσσας Προτροπής	26
3.4.4. Παιδαγωγική Λογική και Σωκρατικός Διάλογος	27
3.4.5. Ιστορικό Συνομιλίας και το Prompt ως αντικείμενο έρευνας.....	28
3.5. Η αρχιτεκτονική του Διακομιστή.....	29
Κεφάλαιο 4: Μεθοδολογία Έρευνας και Αξιολόγησης	31
4.1. Μεθοδολογία Αξιολόγησης	31
4.1.1. Μεθοδολογικό Πλαίσιο	31
4.1.2. Ερευνητικά Ερωτήματα	31
4.2. Ερευνητικός Σχεδιασμός.....	32
4.3. Εργαλεία Συλλογής Δεδομένων και Δειγματοληψία	32
4.3.1. Ερωτηματολόγια	32
4.3.2. Σώμα Κειμένων.....	34
4.3.3. Δειγματοληψία	35
4.4. Ανάλυση Δεδομένων	36
4.4.1. Ποσοτική Ανάλυση	36
4.4.2. Ποιοτική Ανάλυση	37
4.4.3. Τριγωνισμός Δεδομένων	37
Κεφάλαιο 5: Αποτελέσματα Αξιολόγησης.....	38
5.1. Προφίλ Χρηστών.....	38
5.2. Αποτελέσματα Επιπέδων A1-A2	41
5.2.1. Ανάλυση των Αλληλεπιδράσεων με τα Γλωσσικά Μοντέλα.....	41
5.2.2. Αποτελέσματα Ερωτηματολογίου.....	43
5.2.3. Ανοιχτές Ερωτήσεις	45
5.2.4. Συνολική εικόνα για τα επίπεδα A1-A2.....	46
5.3. Αποτελέσματα επιπέδων C1-C2	46
5.3.1. Ανάλυση των Αλληλεπιδράσεων με τα Γλωσσικά Μοντέλα.....	46
5.3.2. Αποτελέσματα Ερωτηματολογίου.....	49

5.3.3. Ανοιχτές ερωτήσεις	51
5.3.4. Συνολική εικόνα C1-C2	51
5.4. Αποτελέσματα Επιπέδων B1-B2	52
5.4.1. Ανάλυση των Αλληλεπιδράσεων με τα Γλωσσικά Μοντέλα.....	52
5.4.2. Αποτελέσματα Ερωτηματολογίου.....	53
5.4.3. Ανοιχτές Ερωτήσεις	55
5.4.4. Συνολική Εικόνα B1-B2	56
5.5. Αποτελέσματα Αξιολόγησης των Εκπαιδευτικών.....	57
5.5.1. Ποσοτική Ανάλυση Ερωτηματολογίου	57
5.5.2. Ποιοτική Ανάλυση — Ανοιχτές Ερωτήσεις	59
5.5.3. Συνολική Εικόνα από την Αξιολόγηση των Εκπαιδευτικών	60
Κεφάλαιο 6: Συζήτηση των ευρημάτων	62
6.1. Η TN ως Γνωστικό Εργαλείο.....	62
6.2. Παιδαγωγική Αποτελεσματικότητα και Χρηστικότητα	63
6.3. Σωκρατικός Διάλογος, Σκαλωσιά και Σταδιακή Απόσυρση	65
Κεφάλαιο 7: Συμπεράσματα και Μελλοντικές Επεκτάσεις.....	68
Βιβλιογραφία.....	70
Παράρτημα I	75
Παράρτημα II	77
Παράρτημα III	81
Παράρτημα IV	86

Εικόνες

Εικόνα 1: Μοντέλο διαμεσολάβησης μεταξύ Μαθητή-Τεχνητής Νοημοσύνης-Γνώσης. Η TN λειτουργεί ως γνωστικό εργαλείο (cognitive tool) και μηχανισμός υποστήριξης (scaffolding), διευκολύνοντας τη μαθησιακή πορεία του εκπαιδευόμενου προς ανώτερα επίπεδα κατανόησης εντός της Ζώνης Επικείμενης Ανάπτυξης (ZEA). Η διαδικασία είναι δυναμική και αμφίδρομη, ενσωματώνοντας συνεχή ανατροφοδότηση και προσαρμογή.7

Εικόνα 2: Αρχιτεκτονική του Προσαρμοστικού Συστήματος Μάθησης. Ο μαθητής αλληλεπιδρά αμφίδρομα με το σύστημα, το οποίο αποτελείται από τρία μοντέλα: το Μοντέλο Δικτύου Γνώσης, το Μοντέλο Γνωστικού Επιπέδου και το Μοντέλο Προσαρμογής. Η συνεχής αλληλεπίδραση μεταξύ των τριών μοντέλων διασφαλίζει την εξατομίκευση της εκπαιδευτικής εμπειρίας σε πραγματικό χρόνο 11

Εικόνα 3: Τα τρία αλληλένδετα στάδια επεξεργασίας λογοτεχνικού κειμένου από το προτεινόμενο σύστημα: ανάλυση ύφους, παιδαγωγική χαρτογράφηση βάσει CEFR και παραγωγή εκπαιδευτικών δραστηριοτήτων. 15

Εικόνα 4: Βασική αρχιτεκτονική συστήματος που βασίζεται στο μοντέλο Client-Server. Από αριστερά προς τα δεξιά τα τρία επίπεδα στα οποία χωρίζεται το σύστημα.20

Εικόνα 5: Διάγραμμα ακολουθίας που εξηγεί τον τρόπο αλληλεπίδρασης μεταξύ των οντοτήτων του συστήματος. Στο διάγραμμα φαίνεται η ροή των αιτημάτων από τον χρήστη και των αποκρίσεων του συστήματος καθώς και ο ασύγχρονος μηχανισμός καταγραφής των δεδομένων για την αξιολόγηση.22

Εικόνα 6: Διάγραμμα αρθρωτής αρχιτεκτονικής (Modular Architecture) του frontend. Στο διάγραμμα απεικονίζεται ο κεντρικός ρόλος του main.js ως ενωχιστική και ο ιεραρχικός γράφος εξαρτήσεων μεταξύ των επιμέρους λειτουργικών μονάδων της εφαρμογής.23

Εικόνα 7: Εσωτερική αρχιτεκτονική διακομιστή. Στο εσωτερικό του φαίνονται τα δύο endpoints επικοινωνίας με τα γλωσσικά μοντέλα και ο μηχανισμός ασύγχρονης καταγραφής writeToLog() των δεδομένων στο Google Sheets για ερευνητική αξιολόγηση.29

Εικόνα 8: Στιγμιότυπα από τη διαδικασία αξιολόγησης στα φροντιστήρια, όπου οι μαθητές αλληλεπιδρούν με την εφαρμογή μέσω φορητών συσκευών υπό την καθοδήγηση της εκπαιδευτικού.35

Εικόνα 9: Ηλικιακή κατανομή χρηστών του συστήματος.	38
Εικόνα 10: Επίπεδο σπουδών των χρηστών του συστήματος που συμμετείχαν στην αξιολόγηση.....	38
Εικόνα 11: Αποτελέσματα στην ερώτηση για ποιο λόγο ο χρήστης επέλεξε την εκμάθηση της ισπανικής γλώσσας.....	39
Εικόνα 12: Συχνότητα προηγούμενης εμπειρίας στη χρήση συστημάτων TN στην εκμάθηση ξένων γλωσσών από τους συμμετέχοντες. Στην κλίμακα 1-4 το 1 αντιστοιχεί στην απάντηση ποτέ ενώ το 4 στην απάντηση πολύ συχνά.....	40
Εικόνα 13: Αξιολόγηση των χρηστών ως προς την δυνατότητα συνεισφοράς της TN στην εκμάθηση ξένων γλωσσών. Στην κλίμακα 1-5 το 1 αντιστοιχεί στην απάντηση Διαφωνώ Απόλυτα ενώ το 5 στην απάντηση Συμφωνώ Απόλυτα.....	40
Εικόνα 14: Μέσοι όροι και τυπική απόκλιση των απαντήσεων στο ερωτηματολόγιο αξιολόγησης για τα επίπεδα A1-A2. Με πράσινο χρώμα επισημαίνονται οι ερωτήσεις με $M.O. \geq 4.50$ και με κόκκινο οι ερωτήσεις με $M.O. \leq 3.00$	44
Εικόνα 15: Μέσοι όροι και τυπική απόκλιση των απαντήσεων στο ερωτηματολόγιο αξιολόγησης για τα επίπεδα B1-B2. Με πράσινο χρώμα επισημαίνονται οι ερωτήσεις με $M.O. \geq 4.50$ και με κόκκινο οι ερωτήσεις με $M.O. \leq 3.00$	54

Πίνακες

Πίνακας 1: Αυτούσιο απόσπασμα καταγραφής (log) αλληλεπίδρασης μαθητή-Gemini (Επίπεδο A2)	41
Πίνακας 2: Βαθμολογίες Likert στην κλίμακα 1-5 (Διαφωνώ απόλυτα-Συμφωνώ απόλυτα) ανά μαθητή και ερώτηση για τις ερωτήσεις 1-12. Οι μαθητές M1 και M3 επέλεξαν το Gemini (γαλάζια απόχρωση), ενώ οι M2 και M4 το Llama (γκρι απόχρωση). Όπου πράσινη σκίαση στην τελευταία στήλη σημαίνει ότι ο M.O. ≥ 4.50 και κόκκινη σκίαση ότι ο M.O. ≤ 3.00	43
Πίνακας 3: Αυτούσιο απόσπασμα καταγραφής (log) αλληλεπίδρασης μαθητή-Gemini (Επίπεδο C2)	47
Πίνακας 4: Βαθμολογίες Likert στην κλίμακα 1-5 (Διαφωνώ απόλυτα-Συμφωνώ απόλυτα) ανά μαθητή και ερώτηση για τις ερωτήσεις E1-E12. Ο μαθητής M5 επέλεξε το γλωσσικό μοντέλο Llama (γκρι απόχρωση), ενώ ο M6 το Gemini (γαλάζια απόχρωση).	49
Πίνακας 5: Βαθμολογίες Likert στην κλίμακα 1-5 (Διαφωνώ απόλυτα-Συμφωνώ απόλυτα) ανά μαθητή και ερώτηση για τις ερωτήσεις 1-12. Οι μαθητές M7-M9 επέλεξαν το Llama (γκρι απόχρωση), ενώ οι M10-M12 το Gemini (γαλάζια απόχρωση). Όπου πράσινη σκίαση στην τελευταία στήλη σημαίνει ότι ο M.O. ≥ 4.50 και κόκκινη σκίαση ότι ο M.O. ≤ 3.00	53
Πίνακας 6: Απαντήσεις των καθηγητών K1 και K2 ανά τύπο ερώτησης: Ποσοτική αξιολόγηση βάσει 5βαθμης κλίμακας Likert (E1-E11, όπου 1: Διαφωνώ απόλυτα έως 5: Συμφωνώ απόλυτα) και συνολική/ποιοτική εκτίμηση της συνεδρίας (E12-E14, με διαβάθμιση από «Πολύ κακή» έως «Άριστη»). Με γαλάζια απόχρωση αποτυπώνονται οι ερωτήσεις που έλαβαν την ίδια βαθμολογία και από τους 2 εκπαιδευτικούς.	57

Κεφάλαιο 1: Εισαγωγή

1.1. Η Ψηφιακή Μεταμόρφωση της Εκπαίδευσης και η Άνοδος της Τεχνητής Νοημοσύνης

Η σύγχρονη εκπαίδευση διανύει μια περίοδο ψηφιακού μετασχηματισμού, με την Τεχνητή Νοημοσύνη (TN) να βρίσκεται στο επίκεντρο, ως ο πλέον καθοριστικός παράγοντας για την αναπροσαρμογή της διδακτικής διαδικασίας. Στις μέρες μας, συστήματα TN επιτρέπουν στους ανθρώπους να μαθαίνουν με τη βοήθεια ψηφιακών βοηθών, όπως τα bots (Fitria, 2021). Χαρακτηριστικά είναι τα παραδείγματα όπου δημοφιλή ψηφιακά εργαλεία όπως το Khanmigo¹, το Duolingo², το LanguaTalk³ και το TeachBetter.ai⁴ κερδίζουν διαρκώς έδαφος ως ολοκληρωμένοι βοηθοί μάθησης.

Πιο συγκεκριμένα, η TN στην εκπαίδευση (edu-AI) δεν είναι απλώς ένα ψηφιακό εργαλείο, αλλά μια θεμελιώδης δομική παρέμβαση, ένας διαμεσολαβητής μεταξύ ανθρώπου και γνώσης, που επηρεάζει την μαθησιακή διαδικασία σε πολλαπλά επίπεδα (Gibson, et al., 2023). Το εύρος αυτών των συστημάτων είναι εντυπωσιακό, καθώς εκτείνονται από το μικρο-επίπεδο της εξατομικευμένης μάθησης (individual learning) μέχρι το μακρο-επίπεδο του θεσμικού μετασχηματισμού ολόκληρου του εκπαιδευτικού συστήματος (π.χ. νέα προγράμματα σπουδών και εκπαιδευτικές πολιτικές που χαράσσονται σε Σχολικό και Πανεπιστημιακό επίπεδο) (Gibson et al., 2023).

Η βασικότερη επιδίωξη της εξατομικευμένης μάθησης είναι η μετάβαση από το ενιαίο, δασκαλοκεντρικό μοντέλο (one style fits all), στο μοντέλο της εξατομικευμένης διδασκαλίας. Στις μέρες μας, η Προσαρμοστική Ηλεκτρονική Μάθηση (adaptive e-learning) αναδεικνύεται ως ο βασικός μηχανισμός που με ηλεκτρονικά μέσα καθιστά την εξατομίκευση αποτελεσματική και βιώσιμη, ενισχύοντας μέσω κινητροδότησης, την ενεργό συμμετοχή των μαθητών στην εκπαιδευτική διαδικασία (El-Sabagh, 2021).

¹ <https://khanmigo.ai/>

² <https://blog.duolingo.com/>

³ <https://languatalk.com/langua>

⁴ <https://teachbetter.ai/>

1.2. Σκοπός της εργασίας

Με γνώμονα τις δυνατότητες που προσφέρει η ΤΝ στην εκπαίδευση, η παρούσα εργασία εστιάζει στη δημιουργία ενός καινοτόμου εργαλείου για τη διδασκαλία της ισπανικής γλώσσας. Μέσω της χρήσης Μεγάλων Γλωσσικών Μοντέλων και της Επεξεργασίας Φυσικής Γλώσσας, επιχειρείται η αυτόματη παραγωγή εκπαιδευτικών δραστηριοτήτων βασισμένων σε λογοτεχνικά κείμενα, με σκοπό την ενίσχυση της εξατομικευμένης μάθησης με προσαρμοσμένες ασκήσεις σύμφωνα με τα επίπεδα γλωσσομάθειας που ορίζει το Κοινό Ευρωπαϊκό Πλαίσιο Αναφοράς (CEFR, A1–C2)⁵. Το προτεινόμενο σύστημα, μια ψηφιακή εκπαιδευτική εφαρμογή, υποστηρίζει τη διδασκαλία της ισπανικής γλώσσας και λογοτεχνίας, αξιοποιώντας τεκμηριωμένες αρχές προσαρμοστικής μάθησης και ευφυή εκπαιδευτικά συστήματα που βασίζονται σε Μεγάλα Γλωσσικά Μοντέλα όπως η οικογένεια GPT⁶, το Claude ⁷(Anthropic) και το Gemini⁸.

Η εργασία οριοθετεί το πεδίο της έρευνας στην τομή τριών βασικών επιστημονικών αξόνων: α) τον Κονστрукτιβισμό, β) την Προσαρμοστική Μάθηση (Adaptive Learning) και γ) την Τεχνητή Νοημοσύνη στην Εκπαίδευση (Edu-AI), με ειδική εφαρμογή στην διδασκαλία της ισπανικής γλώσσας μέσα από γνωστά λογοτεχνικά κείμενα. Η έμφαση δίνεται στον τρόπο με τον οποίο ο παιδαγωγικός σχεδιασμός ενός συστήματος ΤΝ μπορεί να ενισχύσει τις δεξιότητες ανώτερου επιπέδου, όπως η κριτική σκέψη και η δημιουργικότητα, και όχι μόνο η γνωστική απόδοση (Almulla, 2023). Έτσι, η ενσωμάτωση της ΤΝ σε ένα κονστрукτιβιστικό πλαίσιο σηματοδοτεί μια αλλαγή του εκπαιδευτικού παραδείγματος δημιουργώντας δυναμικά και προσαρμοστικά μαθησιακά περιβάλλοντα εκπαίδευσης που προάγουν τη βαθύτερη κατανόηση από τον μαθητή μέσω της αλληλεπίδρασης (Rasul et al., 2023).

1.3. Διάρθρωση της εργασίας

Στην παρούσα ενότητα παρουσιάζεται η διάρθρωση της εργασίας, με συνοπτική αναφορά στα επιμέρους μέρη και το περιεχόμενό τους, ώστε να αποτυπωθεί με

⁵ <https://www.coe.int/en/web/common-european-framework-reference-languages/level-descriptions>

⁶ <https://openai.com/el-GR/index/hello-gpt-4o/>

⁷ <https://www.anthropic.com/news/claude-3-5-sonnet>

⁸ <https://blog.google/innovation-and-ai/products/google-gemini-next-generation-model-february-2024/#sundar-note>

σαφήνεια η δομή και η λογική ανάπτυξη του θέματος. Στο κεφάλαιο 2 επιχειρείται μια θεωρητική μελέτη και σύνθεση της τρέχουσας ερευνητικής στάθμης. Στο κεφάλαιο 3 προτείνεται η αρχιτεκτονική συστήματος. Στο κεφάλαιο 4 παρουσιάζεται η διαδικασία που ακολουθήθηκε κατά την αξιολόγηση του συστήματος. Στο κεφάλαιο 5 συζητούνται τα αποτελέσματα που προέκυψαν από την διαδικασία αξιολόγησης της εφαρμογής και το κεφάλαιο 6 κλείνει με μια συζήτηση για τα συμπεράσματα που προέκυψαν. Τέλος το κεφάλαιο 7 κλείνει με κάποιες μελλοντικές επεκτάσεις και εφαρμογές στο πεδίο.

Κεφάλαιο 2: Θεωρητικό Πλαίσιο και Σύνθεση

2.1.Οριοθέτηση της Ανασκόπησης

Στο πλαίσιο της παρούσας εργασίας, η βιβλιογραφική ανασκόπηση αποσκοπεί στη θεωρητική διερεύνηση και σύνθεση της σύγχρονης ερευνητικής γνώσης, με στόχο τη διαμόρφωση του θεωρητικού και μεθοδολογικού της πλαισίου. Ειδικότερα, επιδιώκεται:

- i. Να αναλυθεί ο Κονστρουκτιβισμός ως το κύριο παιδαγωγικό παράδειγμα για την ανάπτυξη της κριτικής σκέψης και της επίλυσης μη-δομημένων προβλημάτων (ill-structured problems) (Loyens & Gijbels, 2008).
- ii. Να διερευνηθούν οι μηχανισμοί της προσαρμοστικής μάθησης, με έμφαση στην ενεργό συμμετοχή του μαθητή, στα μαθησιακά στυλ και στα κίνητρα ως παράγοντες προσαρμογής (El-Sabagh, 2021; Bayounes et al., 2022).
- iii. Να παρουσιαστεί ο ρόλος της TN στην εκπαίδευση και να εξεταστεί η παιδαγωγικά καθοδηγούμενη ενσωμάτωση των μεγάλων γλωσσικών μοντέλων στα σύγχρονα εκπαιδευτικά περιβάλλοντα (Gibson et al., 2023; Yanthi et al., 2024).
- iv. Τέλος, να εντοπιστεί το ερευνητικό κενό που αφορά την εφαρμογή κονστρουκτιβιστικών και προσαρμοστικών συστημάτων TN στη διδασκαλία λογοτεχνικού περιεχομένου.

2.2. Οι Θεμελιώδεις Αρχές του Κονστρουκτιβισμού

Ο Κονστρουκτιβισμός έχει καθιερωθεί ως η βασική θεωρία μάθησης που στηρίζει τις μεταρρυθμίσεις στη σύγχρονη εκπαίδευση, όπως για παράδειγμα η διερευνητική μάθηση (Stoeckel, 2020). Η κεντρική ιδέα του Κονστρουκτιβισμού είναι ότι οι μαθητές δεν είναι παθητικοί δέκτες πληροφοριών, αλλά αποκλειστικοί δημιουργοί/δομητές της γνώσης. Σύμφωνα με αυτή την ιδέα, η μάθηση, η οποία βασίζεται σε προηγούμενες εμπειρίες, είναι μια ενεργή και διαρκής διαδικασία οικοδόμησης της γνώσης. (Yilmaz, 2008; Saarsar, 2018).

Οι βασικές αρχές που διέπουν ένα κονστρουκτιβιστικό μαθησιακό περιβάλλον περιλαμβάνουν (Saarsar, 2018) την:

- **Ενεργό Συμμετοχή (Active Engagement):** Η μάθηση είναι πιο αποτελεσματική όταν οι μαθητές συμμετέχουν ενεργά στη διαδικασία, δηλαδή ενθαρρύνονται να εξερευνούν, να θέτουν ερωτήματα και να διερευνούν τις απαντήσεις αντί να απομνημονεύουν απλώς δεδομένα (Stoeckel, 2020).
- **Κοινωνική Αλληλεπίδραση (Social Interaction):** Η μάθηση αναπτύσσεται μέσα από κοινωνικές διεργασίες όπως ο διάλογος και η αλληλεπίδραση με άτομα του ευρύτερου περιβάλλοντος, όπως οι συμμαθητές και οι εκπαιδευτικοί (Cole & Scribner, 1978). Σε αυτό το πλαίσιο, τα τελευταία χρόνια, η ΤΝ εμφανίστηκε ως ένας νέος -ψηφιακός- διαμεσολαβητής που αλληλοεπιδρά με τον μαθητή.
- **Γνωστική Σύγκρουση (Cognitive Conflict and Disequilibrium):** Η μάθηση συμβαίνει συχνά όταν οι μαθητές έρχονται αντιμέτωποι με πληροφορίες που έρχονται σε σύγκρουση με την προηγούμενη γνώση τους και τις προϋπάρχουσες πεποιθήσεις τους. Αυτή η γνωστική σύγκρουση οδηγεί τον μαθητή σε αναδόμηση και ανασυγκρότηση της γνώσης και σε «βαθύτερη» μάθηση (Saarsar, 2018).

Ειδικότερα, στη γλωσσική εκπαίδευση υπό το πρίσμα της λογοτεχνίας, ο Κονστρουκτιβισμός δεν σημαίνει απλώς εκμάθηση λεξιλογίου και αποστήθιση γραμματικών κανόνων από τον μαθητή. Μεταφράζεται σε διδακτική προσέγγιση που ενθαρρύνει τον μαθητή να κατανοήσει, να ερμηνεύσει και να εμβαθύνει σε ένα λογοτεχνικό κείμενο, μέσα από γλωσσικές δραστηριότητες. Έτσι η εκμάθηση της γλώσσας έρχεται ως αποτέλεσμα μιας ενεργού εμπλοκής του μαθητή με αυθεντικά λογοτεχνικά ερεθίσματα.

2.2.1. Κριτική Σκέψη και Δημιουργικότητα (Ανάπτυξη Δεξιοτήτων Ανώτερου Επιπέδου)

Η ερευνητική κοινότητα τονίζει ότι ο Κονστρουκτιβισμός αποτελεί το απαραίτητο παράδειγμα για την ανάπτυξη δεξιοτήτων ανώτερου επιπέδου, όπως η κριτική σκέψη (critical thinking), η δημιουργικότητα και η επίλυση προβλημάτων στην εκπαίδευση (Almulla, 2023).

Η ανάπτυξη αυτών των δεξιοτήτων, από τη μια διευκολύνεται από μεθόδους που απαιτούν από τον μαθητή να σκέφτεται κριτικά, να αξιολογεί πληροφορίες και να λαμβάνει τεκμηριωμένες αποφάσεις και από την άλλη, ενθαρρύνει τη δημιουργία

φρέσκων ιδεών και την ανάπτυξη λύσεων που χαρακτηρίζονται από πρωτοτυπία και πρακτικότητα.

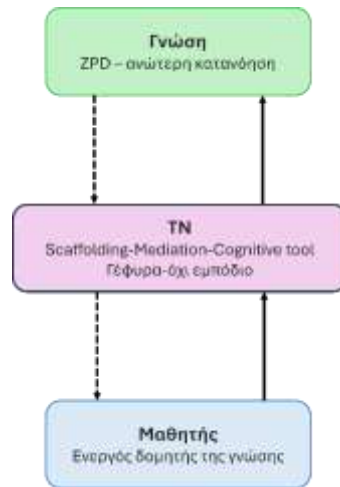
Ειδικότερα, η επίλυση πολύπλοκων μη-δομημένων προβλημάτων (complex ill structured problems) (Loyens & Gijbels, 2008), τα οποία είναι σύνθετα και δεν έχουν μια σαφή, μονοσήμαντη λύση, θεωρείται κεντρική για την ανάπτυξη της κριτικής σκέψης. Η ΤΝ, στο πλαίσιο της προτεινόμενης εργασίας, καλείται να λειτουργήσει ως εργαλείο που διευκολύνει την πρόσβαση σε πληροφορίες και μέσω ανατροφοδότησης επιτρέπει στον μαθητή να επικεντρωθεί στη κατανόηση της γλώσσας και την ερμηνεία των λογοτεχνικών αποσπασμάτων που μελετά.

2.2.2. Η Κοινωνικοπολιτισμική Θεωρία του Vygotsky και τα Γνωστικά Εργαλεία

Η Κοινωνικοπολιτισμική Θεωρία του Vygotsky τονίζει ότι η μάθηση δεν αποτελεί ατομική διεργασία αλλά αναπτύσσεται μέσα από την κοινωνική και την πολιτισμική αλληλεπίδραση του ατόμου. Στο σύγχρονο ψηφιακό περιβάλλον της εκπαίδευσης, όλες οι ψηφιακές τεχνολογίες, συμπεριλαμβανομένης της ΤΝ, λειτουργούν ως γνωστικά εργαλεία που διαμεσολαβούν στη διαδικασία της μάθησης με σκοπό να επεκτείνουν τις δυνατότητες του μαθητή (Gibson et al., 2023; Fadeev, 2019). Έτσι, η κεντρική έννοια της διαμεσολάβησης στην θεωρία του Vygotsky, βρίσκει άμεση εφαρμογή στην ψηφιακή εποχή γιατί τα παραδοσιακά εργαλεία της νόησης, με κυρίαρχο τη γλώσσα, μετεξελίσσονται μέσα από τα σύγχρονα ηλεκτρονικά μέσα και τα νέα ψηφιακά εργαλεία (Fadeev, 2019).

Η παρούσα εργασία επιχειρεί να απαντήσει πώς η ΤΝ λειτουργεί ως το σύγχρονο διαμεσολαβητικό εργαλείο που επιτρέπει στον μαθητή να προσεγγίσει την ισπανική γλώσσα όχι ως στατική πληροφορία, αλλά ως ένα δυναμικό πεδίο αλληλεπίδρασης. Η αξιοποίηση των μεγάλων γλωσσικών μοντέλων για την εμβάθυνση σε διάφορες γλωσσικές και λογοτεχνικές παραμέτρους —όπως το τυπικό και η σύνταξη μιας γλώσσας, η ανάλυση χαρακτήρων, το κοινωνικό και το ιστορικό υπόβαθρο— παρέχει στον μαθητή τις βάσεις, λειτουργώντας ως ένας μηχανισμός γνωστικής πλαισίωσης, γνωστό με τον όρο σκαλωσιά (scaffolding). Με αυτή την διαμεσολάβηση, ο μαθητής δραστηριοποιείται εντός της Ζώνης Επικείμενης Ανάπτυξης (ΖΕΑ), καθώς το σύστημα τον βοηθά να γεφυρώσει το χάσμα ανάμεσα σε όσα ήδη κατέχει και σε ανώτερους

μαθησιακούς στόχους (βλέπε Εικόνα 1) που χωρίς την υποστήριξη ενός έξυπνου συστήματος θα ήταν δύσκολο να επιτευχθούν (Margolis, 2020; Verenikina, 2003).



Εικόνα 1: Μοντέλο διαμεσολάβησης μεταξύ Μαθητή-Τεχνητής Νοημοσύνης-Γνώσης. Η TN λειτουργεί ως γνωστικό εργαλείο (*cognitive tool*) και μηχανισμός υποστήριξης (*scaffolding*), διευκολύνοντας τη μαθησιακή πορεία του εκπαιδευόμενου προς ανώτερα επίπεδα κατανόησης εντός της Ζώνης Επικείμενης Ανάπτυξης (ZEA). Η διαδικασία είναι δυναμική και αμφίδρομη, ενσωματώνοντας συνεχή ανατροφοδότηση και προσαρμογή.

Στο πλαίσιο αυτό, ο ρόλος των μεγάλων γλωσσικών μοντέλων και των αλγορίθμων TN ως «ο ικανότερος άλλος» (More Knowledgeable Other) μετατοπίζεται από την απλή παροχή πληροφοριών στη δυναμική υποστήριξη της οικοδόμησης της γνώσης. Ο «ικανότερος άλλος» δεν είναι μια αυθεντία που κατέχει την αλήθεια, αλλά ένας καταλύτης που ενεργοποιεί τη σκέψη του μαθητή πάνω στο ισπανικό κείμενο. Για να επιτευχθεί αυτό, η ψηφιακή διαμεσολάβηση πρέπει να σχεδιαστεί ως ένας μηχανισμός σταδιακής απόσυρσης της υποστήριξης (*fading*), όπου το σύστημα παρέχει αρχικά έντονη καθοδήγηση —μέσω επεξηγηματικών υποδείξεων και ανατροφοδότησης— η οποία βαθμιαία μειώνεται καθώς ο μαθητής εσωτερικεύει τις στρατηγικές ανάλυσης των κειμένων (Verenikina, 2003).

Συνεπώς, η TN στην παρούσα εργασία δεν αντιμετωπίζεται ως υποκατάστατο της κριτικής σκέψης, αλλά ως ένας έξυπνος συνεργάτης/συνεργός (*co-agent*) που ενισχύει τις γνωστικές δυνατότητες του εκπαιδευόμενου εντός της ZEA (Gibson et al., 2023). Η προσέγγιση αυτή διασφαλίζει ότι η τεχνολογική παρέμβαση παραμένει πιστή στην κοινωνικοπολιτισμική φύση της μάθησης, προάγοντας την αυτονομία του μαθητή. Πρακτικά, στην περίπτωση μας, το σύστημα πρέπει να προσφέρει στον μαθητή

καθοδήγηση πώς να εργαστεί με μια δραστηριότητα, χωρίς να του προσφέρει ξερά την απάντηση.

2.2.3. Ενεργός Δόμηση της Γνώσης και τα Μη-Δομημένα Προβλήματα

Η αποτελεσματικότητα οποιουδήποτε ψηφιακού συστήματος ως μαθησιακό περιβάλλον, δεν αποτελεί κατ' ανάγκη ζήτημα τεχνικής αρτιότητας, αλλά εξαρτάται πρωτίστως από την υποκείμενη παιδαγωγική θεωρία που το διέπει (Gibson et al., 2023). Η θεωρία του κονστρουκτιβισμού προσφέρει το απαραίτητο θεωρητικό υπόβαθρο για τη μετάβαση από την παθητική παροχή πληροφοριών στην ενεργό κινητροδότηση του μαθητή. Σύμφωνα με τον Yilmaz (2008) και όπως αναφέρεται στην ενότητα 2.2, η γνώση δεν αποτελεί ένα έτοιμο προϊόν προς μεταφορά, αλλά μια δυναμική διαδικασία όπου οι μαθητές οικοδομούν νέα νοήματα (meanings) ερμηνεύοντας τις εμπειρίες τους μέσα από το πρίσμα της προϋπάρχουσας γνώσης τους.

Οι Loyens & Gijbels (2008) και Saarsar (2018) αναφέρουν ότι τα Κονστρουκτιβιστικά Μαθησιακά Περιβάλλοντα (Constructivist Learning Environments - CLEs) διαφοροποιούνται ριζικά από τα παραδοσιακά μοντέλα διδασκαλίας μέσω συγκεκριμένων αξόνων που προωθούν μια βαθύτερη γνωστική εμπλοκή. Κεντρικό σημείο αυτής της διαφοροποίησης αποτελεί η μαθητοκεντρική προσέγγιση, όπου ο μαθητής αναλαμβάνει την πλήρη ευθύνη της πορείας του. Ο μαθητής πρέπει να λειτουργεί αυτόνομα και με δική του πρωτοβουλία να δράσει, θέτοντας ερωτήματα και επιλύοντας σύνθετα προβλήματα (Saarsar, 2018).

Παράλληλα, τα κονστρουκτιβιστικά περιβάλλοντα ενσωματώνουν πολύπλευρες αναπαραστάσεις της γνώσης. Για παράδειγμα, το ίδιο απόσπασμα μπορεί να προσεγγιστεί γλωσσολογικά, αφηγηματολογικά ή ιστορικά — και κάθε οπτική αποκαλύπτει κάτι διαφορετικό. Με την στρατηγική αυτή ο μαθητής προσεγγίζει τη γνώση από διαφορετικές οπτικές γωνίες, αποφεύγοντας παράλληλα την υπεραπλούστευση (Loyens & Gijbels, 2008). Η μαθησιακή διαδικασία μετατρέπεται έτσι σε μια συνεχή εμπειρία ενεργού εξερεύνησης και πειραματισμού, στην οποία ο εκπαιδευτικός παύει να είναι η μοναδική πηγή αυθεντίας και λειτουργεί πλέον ως καθοδηγητής και αρωγός (facilitator), ενθαρρύνοντας τον μαθητή να αλληλεπιδράσει ουσιαστικά με το υλικό (Stoeckel, 2020). Ο μαθητής όταν διαβάζει ένα κείμενο δεν πρέπει απλά να το αναπαράγει αλλά να το αμφισβητεί, κάτι που απαιτεί από το

σύστημα να του θέτει τα κατάλληλα ερωτήματα οδηγώντας τον προς την κριτική αμφισβήτηση και όχι να του προσφέρει έτοιμες τις απαντήσεις.

Επιπλέον, η γνωστική ανάπτυξη ενισχύεται σημαντικά από το στοιχείο της αυθεντικότητας, καθώς η μάθηση τοποθετείται μέσα σε πραγματικά πλαίσια και συνδέεται με προβλήματα της αληθινής ζωής που έχουν ουσιαστικό νόημα για τον μαθητή (Loyens & Gijbels, 2008). Αυτή η σύνδεση της θεωρίας με την πράξη εξασφαλίζει ότι η γνώση δεν παραμένει στατική, αλλά οικοδομείται ενεργά μέσα από την εμπειρία - ιδιαίτερα στη μελέτη της γλώσσας, όπου η επαφή με αυθεντικά λογοτεχνικά κείμενα προσδίδει νόημα στη μαθησιακή προσπάθεια. Γι' αυτό και η χρήση νέων τεχνολογιών μπορεί να υποστηρίξει την οικοδόμηση της γνώσης, προσφέροντας εξατομικευμένα μονοπάτια που ανταποκρίνονται στις ιδιαίτερες γνωστικές ικανότητες και τις μαθησιακές συνήθειες του κάθε μαθητή.

Η αξία του Κονστрукτιβισμού γίνεται ιδιαίτερα εμφανής όταν ο μαθητής καλείται να διαχειριστεί τα λεγόμενα μη-δομημένα προβλήματα (ill-structured problems). Σε αντίθεση με τα καλά-δομημένα προβλήματα (well-structured problems), τα οποία διαθέτουν σαφείς στόχους και προβλέψιμες λύσεις (τυποποιημένες ασκήσεις), τα μη-δομημένα προβλήματα χαρακτηρίζονται από ασαφείς στόχους, ελλιπή δεδομένα και την απουσία μιας μοναδικής ορθής απάντησης (Voss & Post, 1988, όπως αναφέρεται στο Loyens & Gijbels, 2008; Jonassen, 1997). Η αντιμετώπισή τους προϋποθέτει από τον μαθητή να αξιολογήσει πολλαπλές ερμηνευτικές εκδοχές, να σταθμίσει αντικρουόμενα επιχειρήματα και να καταλήξει σε μια δική του, τεκμηριωμένη και αιτιολογημένη θέση (Almulla, 2023; Loyens & Gijbels, 2008).

Κατά συνέπεια εδώ, η γλωσσική και λογοτεχνική προσέγγιση αποσπασμάτων από έργα της ισπανόφωνης γραμματείας και η παραγωγή ασκήσεων μέσα από αυτά, αντιμετωπίζεται ως ένα κατεξοχήν μη-δομημένο πρόβλημα. Καθώς δεν υπάρχει μία αντικειμενική αλήθεια για την ερμηνεία ενός κειμένου, η χρήση της ΤΝ ως γνωστικού εργαλείου και καθοδηγητή, στοχεύει στην παροχή ενός προσαρμοστικού πλαισίου υποστήριξης-σκαλωσιά. Αυτό επιτρέπει στον μαθητή να δραστηριοποιείται εντός της ΖΕΑ, επιτυγχάνοντας στόχους που υπερβαίνουν το επίπεδο της αυτόνομης επίδοσής του (Margolis, 2020; Verenikina, 2003). Με άλλα λόγια, η ΤΝ δεν καλείται να δώσει στον μαθητή τις απαντήσεις, αλλά να τον βοηθήσει να φτάσει σε σημείο που να μπορεί να τις διατυπώσει μόνος του.

2.3. Προσαρμοστική Μάθηση

2.3.1. Ορισμός και Μηχανισμοί της Προσαρμοστικής Μάθησης (Adaptive Learning)

Η Προσαρμοστική Μάθηση (Adaptive Learning) ορίζεται ως ένα «έξυπνο» σύστημα εξατομικευμένης διδασκαλίας που τροποποιεί δυναμικά το περιεχόμενο, τη σειρά και την δομή των εκπαιδευτικών δραστηριοτήτων. Στόχος της προσαρμοστικής μάθησης είναι να παρεμβαίνει υποστηρικτικά έτσι ώστε να προσαρμόζεται στις πραγματικές ανάγκες, τις προτιμήσεις και την πρόοδο του κάθε μαθητή (Fadieieva, 2023). Αντίθετα με τα συμβατικά ηλεκτρονικά συστήματα εκπαίδευσης, η προσαρμοστική μάθηση λειτουργεί ως ένας μηχανισμός κινητροδότησης που αποσκοπεί στη βελτίωση των μαθησιακών αποτελεσμάτων μέσα από μια διαδικασία που είναι γνωστή ως μαθησιακή εμπλοκή (student engagement). Συνεπώς, ο βασικός στόχος της Προσαρμοστικής Μάθησης είναι η βελτιστοποίηση της διδακτικής εμπειρίας (El-Sabagh, 2021).

Η λειτουργία τέτοιων συστημάτων βασίζεται σε τρεις κύριες ενότητες και στη συνεχή αλληλεπίδραση μεταξύ τους (βλέπε Εικόνα 2): το μοντέλο δικτύου γνώσης (knowledge network model), το μοντέλο γνωστικού επιπέδου (cognitive level model) και το μοντέλο προσαρμογής (adaptive recommendation model). Το μοντέλο δικτύου γνώσης αποτελεί τον πυρήνα του συστήματος, καθώς καταγράφει δυναμικά το γνωστικό προφίλ, το μαθησιακό στυλ και τις αντιδράσεις του μαθητή, επιτρέποντας στο σύστημα να προσαρμόζεται σε πραγματικό χρόνο βελτιώνοντας έτσι την εκπαιδευτική εμπειρία (El-Sabagh, 2021; Zhao & Wang, 2019). Το μοντέλο του γνωστικού επιπέδου οργανώνει το διδακτικό υλικό σε ενότητες, σχεδιασμένες ώστε να ανταποκρίνονται στο επίπεδο και στις προτιμήσεις του εκάστοτε μαθητή (El-Sabagh, 2021; Fadieieva, 2023). Τέλος, το μοντέλο προσαρμογής αξιοποιεί μηχανισμούς — όπως αλγορίθμους μηχανικής μάθησης — για να γεφυρώσει τις ανάγκες του μαθητή με το κατάλληλο υλικό, διασφαλίζοντας την ακεραιότητα της μαθησιακής διαδικασίας (Zhao & Wang, 2019; El-Sabagh, 2021). Ειδικότερα σε αυτό το επίπεδο, χρησιμοποιείται μια τεχνική, γνωστή ως *reinforcement learning*, δηλαδή ένα σύστημα που μαθαίνει δυναμικά από τις απαντήσεις του μαθητή και προσαρμόζεται, όπως ακριβώς και ένας προσωπικός δάσκαλος προσαρμόζεται συνεχώς στις ανάγκες και το επίπεδο του μαθητή (Fahad et al., 2023).



***Εικόνα 2:** Αρχιτεκτονική του Προσαρμοστικού Συστήματος Μάθησης. Ο μαθητής αλληλεπιδρά αμφίδρομα με το σύστημα, το οποίο αποτελείται από τρία μοντέλα: το Μοντέλο Δικτύου Γνώσης, το Μοντέλο Γνωστικού Επιπέδου και το Μοντέλο Προσαρμογής. Η συνεχής αλληλεπίδραση μεταξύ των τριών μοντέλων διασφαλίζει την εξατομίκευση της εκπαιδευτικής εμπειρίας σε πραγματικό χρόνο*

2.3.2. Μαθησιακά Στυλ και Σχεδιασμός Κινήτρων (Μοντέλο ARCS)

Η ενσωμάτωση των μαθησιακών στυλ στα προσαρμοστικά περιβάλλοντα θεωρείται θεμελιώδης γιατί δεν μαθαίνουμε όλοι με τον ίδιο τρόπο. Σύμφωνα με τον El-Sabagh (2021), όταν οι ψηφιακές μέθοδοι και οι διδακτικές στρατηγικές ταιριάζουν με τον τρόπο που ο καθένας μαθαίνει, η κατανόηση και η αφομοίωση της γνώσης βελτιώνονται σημαντικά. Οι Havard et al. (2016) προτείνουν μια κρίσιμη διάκριση μεταξύ των προσαρμοστικών συστημάτων, δηλαδή τα συστήματα που αποφασίζουν μόνα τους και προσαρμόζουν το περιεχόμενο και των παραμετροποιήσιμων, όπου ο ίδιος ο μαθητής μπορεί να διαμορφώσει το περιβάλλον σύμφωνα με τις ανάγκες του. Αυτή η διάκριση από την αμιγώς τεχνολογική αυτοματοποίηση της διαδικασίας στην ενεργό εμπλοκή του μαθητή αντικατοπτρίζει μια ουσιαστική αλλαγή φιλοσοφίας. Την μετάβαση από την ιδέα ότι το σύστημα αποφασίζει για τον μαθητή, στην ιδέα ότι ο ίδιος ο μαθητής είναι αυτός που διαμορφώνει την μαθησιακή του εμπειρία.

Ενώ τα μαθησιακά στυλ καθορίζουν τον τρόπο πρόσληψης της πληροφορίας, η επιτυχία ενός προσαρμοστικού συστήματος βασίζεται εξίσου στη διατήρηση των κινήτρων του μαθητή. Το μοντέλο ARCS του Keller (1987) παρέχει το πλαίσιο για τη δυναμική ενίσχυση της εμπλοκής (engagement) μέσω TN. Έτσι, μέσω της χρήσης της TN, το σύστημα μπορεί να διεγείρει την περιέργεια (Attention), να συνδέει το περιεχόμενο με τους προσωπικούς στόχους του μαθητή (Relevance), να παρέχει κατάλληλη υποστήριξη που οικοδομεί την αυτοπεποίθηση (Confidence) και να προσφέρει ποιοτική ανατροφοδότηση που οδηγεί στην ικανοποίηση (Satisfaction). Στην περίπτωση μας, δηλαδή την εκμάθηση μιας ξένης γλώσσας, είναι σημαντικό το

σύστημα να μην επιλέγει απλώς το περιεχόμενο, αλλά να προσφέρει ασκήσεις που αγγίζουν τα προσωπικά ενδιαφέροντα του μαθητή. Η εφαρμογή πρέπει να δείχνει στον μαθητή ότι τα καταφέρνει μέσα από σταδιακά προκλητικές δραστηριότητες, και να του δίνει ανατροφοδότηση που αφήνει την αίσθηση της προόδου — όχι της αποτυχίας (El-Sabagh, 2021).

2.3.3. Η Μαθησιακή Εμπλοκή και τα Κίνητρα

Η αποτελεσματικότητα των τεχνολογικών παρεμβάσεων στην εκπαίδευση κρίνεται τελικά από το επίπεδο της μαθησιακής εμπλοκής, η οποία αποτελεί πολυδιάστατο δείκτη της ποιότητας της μάθησης. Σύμφωνα με τη σχετική βιβλιογραφία, τα προσαρμοστικά περιβάλλοντα παρουσιάζουν σημαντικά υψηλότερη εμπλοκή σε σύγκριση με τις παραδοσιακές μεθόδους γιατί επιτρέπουν στον μαθητή να κινείται με τους δικούς του ρυθμούς (El-Sabagh, 2021). Σύμφωνα με τους Bayounes et al. (2022), η καθοδήγηση του μαθητή κατά την διάρκεια της μάθησης πρέπει να βασίζεται κατά κύριο λόγο στην πρόθεση του μαθητή και συνδυάζει την διδακτική στρατηγική που ακολουθείται με τις προσωπικές ανάγκες και την αυτονομία του μαθητή.

Η μαθησιακή εμπλοκή είναι μια πολυδιάστατη έννοια που αξιολογείται γύρω από δύο βασικούς άξονες. Ο πρώτος άξονας, ο συναισθηματικός, αφορά στο ενδιαφέρον και την ευχαρίστηση που βιώνει ο μαθητής κατά τη μαθησιακή διαδικασία — αλλά και στη μείωση του άγχους που συχνά προκαλούν υπερβολικά απαιτητικά ή ασαφή ερωτήματα (Fadieieva, 2023). Για παράδειγμα, στο πλαίσιο της εκμάθησης μιας ξένης γλώσσας, αυτό έχει ιδιαίτερη σημασία γιατί όταν ένας μαθητής αγχώνεται με ένα λογοτεχνικό απόσπασμα πάνω από το επίπεδό του, απλώς σταματά να προσπαθεί. Ο δεύτερος, ο συμπεριφορικός άξονας, αφορά την επιμονή στην προσπάθεια και την ενεργό αλληλεπίδραση του μαθητή με το σύστημα. Για να επιτευχθεί αυτό, η Προσαρμοστική Μάθηση πρέπει να κατευθύνει τη διδασκαλία στα δυνατά σημεία του μαθητή και να παρέχει άμεση ανατροφοδότηση (Fadieieva, 2023; El-Sabagh, 2021). Με άλλα λόγια, ένα σύστημα που εμπλέκει πραγματικά τον μαθητή δεν είναι αυτό που δίνει τις σωστές απαντήσεις — είναι αυτό που τον κάνει να θέλει να ξαναπροσπαθήσει.

Συνοψίζοντας, η ενσωμάτωση των αρχών της Προσαρμοστικής Μάθησης αποτελεί τον συνδετικό κρίκο μεταξύ της κονστрукτιβιστικής θεωρίας και της τεχνολογικής υλοποίησης του προτεινόμενου συστήματος. Στην περίπτωση της διδασκαλίας της ισπανόφωνης γλώσσας μέσα από πραγματικά κείμενα από την ισπανόγλωσση

γραμματεία, η προσαρμοστικότητα δεν περιορίζεται στην απλή αυξομείωση της γλωσσικής δυσκολίας, αλλά επεκτείνεται στη διαβάθμιση της ερμηνευτικής πρόκλησης. Το σύστημα, πρέπει να λειτουργεί ως ένας ευφυής διαμεσολαβητής, να αναλύει το προφίλ του μαθητή και να παρέχει εξατομικευμένη σκαλωσιά (El-Sabagh, 2021). Αυτό επιτρέπει σε έναν αρχάριο χρήστη να κατανοήσει τη βασική πλοκή ενός δύσκολου κειμένου, ενώ ταυτόχρονα καθοδηγεί έναν προχωρημένο μαθητή στην επίλυση των γνωστών «μη-δομημένων προβλημάτων» που επιβάλλει η λογοτεχνική ανάλυση (Loyens & Gijbels, 2008). Με αυτόν τον τρόπο, η Τεχνητή Νοημοσύνη μετατρέπεται από ένα στατικό εργαλείο σε ένα δυναμικό περιβάλλον που καλλιεργεί τη μαθησιακή εμπλοκή και την κριτική αυτονομία, διασφαλίζοντας ότι κάθε μαθητής αλληλεπιδρά με την ισπανική γραμματεία στο μέγιστο των δυνατοτήτων του εντός της ZEA.

2.4. Η TN στην Εκπαίδευση (Edu-AI)

2.4.1. Το Μοντέλο Τριών Επιπέδων (Micro, Meso, Macro)

Η ενσωμάτωση της TN στην εκπαίδευση δεν αποτελεί μια απλή προσθήκη νέων ψηφιακών εργαλείων στην εκπαιδευτική διαδικασία, αλλά απαιτεί καθολική παιδαγωγική προσέγγιση. Για την κατανόηση του ρόλου της TN στην εκπαίδευση, οι Gibson et al. (2023) προτείνουν ένα μοντέλο τριών επιπέδων —Micro, Meso και Macro— το οποίο προσεγγίζει την μαθησιακή διαδικασία ως ένα δίκτυο αλληλεπιδράσεων σε διαφορετικές κλίμακες.

Το Micro-επίπεδο εστιάζει στην άμεση αλληλεπίδραση του μαθητή με το σύστημα TN σε πραγματικό χρόνο. Όπως υποστηρίζουν οι Guo, Yi, & Liu (2024), τα συστήματα TN μπορούν να παρέχουν εξατομικευμένη ανατροφοδότηση, ενισχύοντας τη μάθηση σύμφωνα με τις αρχές του κονστρουκτιβισμού. Παράλληλα, οι αλγόριθμοι δρουν ως διαμεσολαβητές που προσφέρουν την απαραίτητη σκαλωσιά ώστε ο μαθητής να δραστηριοποιείται εντός της ZEA (Gibson et al., 2023).

Στο Meso-επίπεδο, η TN επεκτείνεται πέρα από το άτομο και επιδρά στην μαθησιακή διαδικασία σε επίπεδο ομάδας ή τάξης. Σε αυτό το επίπεδο, τα συστήματα TN ενισχύουν τον εκπαιδευτικό σχεδιασμό και συνεισφέρουν στη δημιουργία εκπαιδευτικού υλικού και στην αξιολόγηση των μαθητών (Fitria, 2021). Δηλαδή, η TN συμβάλλει στη δημιουργία «έξυπνου περιεχομένου» και ενισχύει τη μαθησιακή

εμπλοκή μέσα από διαδραστικές δραστηριότητες. Κατά συνέπεια, η μάθηση αντιμετωπίζεται ως μια συλλογική διαδικασία όπου η πληροφορία διαχέεται και ανατροφοδοτείται, ενισχύοντας τη μαθησιακή εμπλοκή (Gibson et al., 2023).

Στο Macro-επίπεδο, η TN εξετάζεται υπό το πρίσμα του ευρύτερου εκπαιδευτικού συστήματος και της κοινωνίας. Σε αυτό το επίπεδο, σύμφωνα με τους Crompton & Burke (2023), ζητήματα όπως η ηθική χρήση της TN, η προστασία δεδομένων και η ισότιμη πρόσβαση στην εκπαίδευση έχουν ιδιαίτερη σημασία. Παράλληλα, αναδεικνύονται σύγχρονες προκλήσεις όπως η αλγοριθμική μεροληψία (bias) και η ανάγκη χάραξης εκπαιδευτικών πολιτικών για την «ορθή» χρήση της TN.

Συνεπώς, η παρούσα εργασία επικεντρώνεται κατά κύριο λόγο στο Micro-επίπεδο, καθώς εξετάζει την αλληλεπίδραση του μαθητή σε πραγματικό χρόνο με ένα σύστημα TN για την εκμάθηση της ισπανικής γλώσσας μέσα από την λογοτεχνία. Παράλληλα, το προτεινόμενο σύστημα επεκτείνεται και στο Meso-επίπεδο, καθώς μπορεί να αξιοποιηθεί από έναν εκπαιδευτικό σε επίπεδο ομάδας για την παραγωγή ασκήσεων από συγκεκριμένα κείμενα που διδάσκονται στην τάξη.

2.4.2. Η Ενσωμάτωση των Μεγάλων Γλωσσικών Μοντέλων (LLMs)

Η εμφάνιση των μεγάλων γλωσσικών μοντέλων, τα οποία εντάσσονται στο ευρύτερο πλαίσιο της Παραγωγικής TN (Generative AI), έχει επιφέρει μια επαναστατική αλλαγή στον τρόπο με τον οποίο η TN προσεγγίζει τη γλώσσα, μετατοπίζοντας το ενδιαφέρον από την απλή ανάκτηση πληροφοριών στην παραγωγή περιεχομένου με βάση τα συμφοραζόμενα (Håkansson et al., 2024). Στη γλωσσική εκπαίδευση μέσω της λογοτεχνίας, η πρόκληση είναι διττή: αφενός η κατανόηση του σύνθετου λογοτεχνικού ύφους και αφετέρου η μετατροπή του σε στοχευμένες εκπαιδευτικές δραστηριότητες.

Τα μεγάλα γλωσσικά μοντέλα βασίζονται σε μεγάλο βαθμό στην επεξεργασία φυσικής γλώσσας, η οποία τους επιτρέπει να αναλύουν τη σημασιολογία και τη συντακτική δομή ενός κειμένου. Όπως επισημαίνει η Fitria (2021), η δυνατότητα της TN να λειτουργεί ως «εικονικός μέντορας» βασίζεται στην ικανότητά της να προσομοιώνει την ανθρώπινη συμπεριφορά (*"behave like humans"*, σελ. 138). Στην περίπτωσή μας, αυτό σημαίνει ότι ένα γλωσσικό μοντέλο πρέπει να αναγνωρίζει όχι μόνο τις λέξεις, αλλά και τις μεταφορές, τους ιδιωτισμούς και το ιδιαίτερο ύφος των δημιουργών. Έτσι, η μετάβαση από το λογοτεχνικό κείμενο στην εκπαιδευτική δραστηριότητα

αποτελεί κρίσιμο στάδιο της διαδικασίας, καθώς εκεί υλοποιείται η μετατροπή της παθητικής ανάγνωσης σε ενεργό μάθηση.

Οι Guo et al. (2024) επισημαίνουν ότι η Γενετική ΤΝ δεν περιορίζεται στην αναπαραγωγή υπάρχουσας γνώσης, αλλά έχει τη δυνατότητα να δημιουργεί εξατομικευμένα μαθησιακά μονοπάτια που ανταποκρίνονται στις ατομικές ανάγκες του εκπαιδευόμενου. Εμπνευσμένο από αυτή τη λογική, το προτεινόμενο σύστημα οργανώνει τη διαδικασία παραγωγής εκπαιδευτικών δραστηριοτήτων σε τρία αλληλένδετα στάδια (βλέπε Εικόνα 3). Ξεκινά από την ανάλυση ύφους, όπου το μοντέλο εντοπίζει τα γλωσσικά χαρακτηριστικά του κειμένου — λεξιλόγιο, χρόνους ρημάτων, βαθμό δυσκολίας κ.λπ. Στη συνέχεια ακολουθεί η παιδαγωγική χαρτογράφηση, κατά την οποία τα γλωσσικά χαρακτηριστικά πρέπει να σχετίζονται με το Κοινό Ευρωπαϊκό Πλαίσιο Αναφοράς, αξιοποιώντας ένα προσαρμοστικό μοντέλο παρόμοιο με αυτό που περιγράφουν οι Bayounes et al. (2022), ώστε το σύστημα να προσαρμόζεται κατάλληλα ανάλογα με το επίπεδο του μαθητή. Η διαδικασία ολοκληρώνεται με την παραγωγή ασκήσεων, όπου το μοντέλο μετατρέπει το λογοτεχνικό ερέθισμα σε εκπαιδευτικές δραστηριότητες.



Εικόνα 3: Τα τρία αλληλένδετα στάδια επεξεργασίας λογοτεχνικού κειμένου από το προτεινόμενο σύστημα: ανάλυση ύφους, παιδαγωγική χαρτογράφηση βάσει CEFR και παραγωγή εκπαιδευτικών δραστηριοτήτων.

Ως προς την αποτελεσματικότητα αυτής της προσέγγισης, οι Yanthi et al. (2024) έδειξαν ότι ένα μοντέλο βασισμένο στο ChatGPT μπορεί να ενισχύσει σημαντικά τις γνωστικές δεξιότητες, επιβεβαιώνοντας ότι η ΤΝ λειτουργεί αποτελεσματικότερα όταν εντάσσεται σε κonstruktivistικό πλαίσιο που ενθαρρύνει τον μαθητή να επεξεργάζεται ενεργά την πληροφορία (Almulla, 2023).

Ωστόσο, στην περίπτωση μας, βασική πρόκληση στην ενσωμάτωση τέτοιων μοντέλων στη μαθησιακή διαδικασία παραμένει η διατήρηση της λογοτεχνικής αξίας του αρχικού κειμένου και η διασφάλιση ότι η μετατροπή του σε εκπαιδευτικό υλικό δεν οδηγεί σε

απώλεια του πολιτισμικού βάθους — ένας στόχος που στην ισπανόφωνη λογοτεχνία αποκτά ιδιαίτερη βαρύτητα, δεδομένης της πολιτισμικής πολυφωνίας που τη χαρακτηρίζει.

2.4.3. Ηθικές Διαστάσεις: Αλγοριθμική Μεροληψία και Πολιτισμική Ευαισθησία

Η ενσωμάτωση των μεγάλων γλωσσικών μοντέλων στην εκπαίδευση δεν είναι πανάκεια και δεν αποτελεί απλώς ένα τεχνικό ή παιδαγωγικό ζήτημα αλλά φέρει μαζί της ένα σύνολο από ηθικές προκλήσεις. Οι Rasul et al. (2023) επισημαίνουν πέντε κεντρικές προκλήσεις στη χρήση του ChatGPT στην τριτοβάθμια εκπαίδευση. Εξέχουσα θέση μεταξύ αυτών κατέχουν η μεροληψία και η παραγωγή παραπλανητικών πληροφοριών, γνωστές και ως ψευδαισθήσεις (hallucinations). Οι συγγραφείς τονίζουν ότι τόσο οι εκπαιδευτικοί όσο και οι φοιτητές οφείλουν να ασκούν κριτική επαλήθευση σε κάθε αποτέλεσμα που παράγεται από το μοντέλο.

Η αλγοριθμική μεροληψία αποτελεί συστημικό χαρακτηριστικό των μεγάλων γλωσσικών μοντέλων. Ο Z. Liu (2023) αναφέρει χαρακτηριστικά ότι η μεροληψία οφείλεται κατά κύριο λόγο σε τρεις παράγοντες: την ποιότητα των δεδομένων εκπαίδευσης, τον σχεδιασμό των αλγορίθμων και τη δυναμική φύση της αλληλεπίδρασης των μοντέλων με τους χρήστες.

Ως προς τον πρώτο παράγοντα, στο άρθρο αναφέρεται ότι αν τα δεδομένα εκπαίδευσης ενός μοντέλου προέρχονται κυρίως από συγγραφείς ανδρικού, λευκού και δυτικού υπόβαθρου, τότε το μοντέλο αναπαράγει αναπόφευκτα τις οπτικές και τα στερεότυπα αυτής της ομάδας, αγνοώντας την ποικιλομορφία. Η διαπίστωση αυτή είναι κρίσιμη για τη διδασκαλία μέσα από κείμενα της ισπανόφωνης λογοτεχνίας γιατί συστήματα που έχουν εκπαιδευτεί κυρίως σε τέτοια σύνολα δεδομένων είναι λογικό να ερμηνεύουν τα ισπανόφωνα-λατινοαμερικάνικα λογοτεχνικά κείμενα μέσα από ένα δυτικοκεντρικό πρίσμα, αλλοιώνοντας έτσι τη βαθύτερη πολιτισμική τους ταυτότητα.

Ο δεύτερος παράγοντας αφορά την εγγενή αδιαφάνεια των σύγχρονων μοντέλων, τα οποία συχνά περιγράφονται ως «μαύρα κουτιά» (Adadi & Berrada, 2018). Σύμφωνα με τον Z. Liu (2023), αυτή η ενθυλάκωση των εσωτερικών λειτουργιών που χαρακτηρίζουν τους αλγορίθμους TN, υπονομεύει την εμπιστοσύνη των χρηστών και συγκρούεται με την απαίτηση για επεξηγησιμότητα. Στο εκπαιδευτικό περιβάλλον, ο κίνδυνος είναι ιδιαίτερα ορατός γιατί όταν ένας μαθητής λαμβάνει μια ερμηνεία ή μια

ανατροφοδότηση χωρίς να κατανοεί τη λογική πίσω από αυτήν, αδυνατεί να την αξιολογήσει κριτικά. Δηλαδή αποδέχεται άκριτα την έξοδο του συστήματος ως αδιαμφισβήτητη αυθεντία.

Ο τρίτος παράγοντας αφορά τους βρόγχους ανατροφοδότησης που αναπτύσσονται κατά την αλληλεπίδραση των χρηστών με τα μοντέλα. Όταν οι χρήστες επιβραβεύουν αποτελέσματα που απλώς καθρεφτίζουν τις δικές τους πολιτισμικές προκαταλήψεις, η TN τα εκλαμβάνει ως προτιμητέα, ενισχύοντας έτσι περαιτέρω τις υπάρχουσες μεροληψίες (Z. Liu, 2023). Μια τέτοια δυναμική προσαρμογή υπονομεύει τη θεμελιώδη προσπάθεια της εκπαίδευσης να καλλιεργήσει την κριτική σκέψη και τον σεβασμό στην πολιτισμική ετερότητα.

Τέλος, οι λεγόμενες ψευδαισθήσεις των γλωσσικών μοντέλων προκύπτουν από τον πιθανοκρατικό τρόπο λειτουργίας τους. Τα συστήματα TN δεν διαθέτουν μηχανισμούς επαλήθευσης της αλήθειας. Αντίθετα, λειτουργούν προβλέποντας ποια είναι η πιθανότερη λέξη που πρέπει να ακολουθήσει την προηγούμενη, ώστε να δημιουργηθεί μια λογική πρόταση. Αυτό έχει ως αποτέλεσμα να παράγουν κείμενα που φαίνονται απόλυτα σωστά και πειστικά στη δομή τους, αλλά μπορεί να περιέχουν στοιχεία που είναι εντελώς φανταστικά ή ανακριβή (Ji et al., 2023).

Όλες οι παραπάνω διαπιστώσεις έχουν συγκεκριμένες συνέπειες που πρέπει να ληφθούν σοβαρά υπόψη κατά τον σχεδιασμό του εκπαιδευτικού συστήματος ώστε να μην αλλοιώνουν την πολιτισμική πολυφωνία της ισπανόφωνης γραμματείας. Κεντρική δικλείδα ασφαλείας σε αυτό το εγχείρημα αποτελεί η «ορθή» μηχανική προτροπών (prompt engineering). Μέσω αυτής, η αλληλεπίδραση με το μοντέλο παύει να είναι τυχαία και αποκτά συγκεκριμένες παραμέτρους που περιορίζουν τις μεροληψίες και τις ψευδαισθήσεις (Mollick & Mollick, 2023). Κλείνοντας, αξίζει να παρατηρήσουμε ότι ο Z. Liu (2023) επισημαίνει ότι η διαχείριση της πολιτισμικής μεροληψίας, δεν είναι αποκλειστικά πρόβλημα αλλά αποτελεί μια ευκαιρία για την ενίσχυση της αυτοεπίγνωσης και την εκτίμηση της διαφορετικότητας.

Κεφάλαιο 3: Αρχιτεκτονική Συστήματος

3.1. Σχεδιαστικές Αρχές

3.1.1. Παιδαγωγική Προσέγγιση της Αρχιτεκτονικής

Το παρόν κεφάλαιο παρουσιάζει την αρχιτεκτονική του συστήματος. Οι τεχνολογίες επιλέχθηκαν με βάση τις παιδαγωγικές αρχές όπως αναλύθηκαν στο Κεφάλαιο 2 της παρούσας εργασίας. Ο αρχιτεκτονικός σχεδιασμός ακολουθεί μια ανθρωποκεντρική προσέγγιση με τον μαθητή να βρίσκεται στο επίκεντρο της μαθησιακής διαδικασίας. Έγινε προσπάθεια η εφαρμογή να δομηθεί ως ένας ψηφιακός βοηθός σκαλωσιάς, που καθοδηγεί τον μαθητή εντός της ΖΕΑ του και όχι ως μια απλή εφαρμογή που ανακτά πληροφορίες. Δηλαδή, το σύστημα δεν αποτελεί ένα απλό περιτύλιγμα γύρω από ένα γλωσσικό μοντέλο, αλλά μια πλήρως λειτουργική εφαρμογή που εκτός του παιδαγωγικού πλαισίου που επιχειρεί να υποστηρίξει, ενσωματώνει πολυτροπικές δυνατότητες όπως η αναγνώριση και σύνθεση ομιλίας και η υποδομή για την ερευνητική καταγραφή δεδομένων σε πραγματικό χρόνο.

3.1.2. Επιλογή Τεχνολογικού Stack

Το σύστημα ακολουθεί την γνωστή αρχιτεκτονική Πελάτη-Εξυπηρετητή (Fielding, 2000), στην οποία το σύστημα στο προσκήνιο (frontend) περιλαμβάνει τη διεπαφή με τον χρήστη και στο παρασκήνιο (backend) λειτουργεί ως διαμεσολαβητής προς τις εξωτερικές υπηρεσίες. Στην επιλογή των εργαλείων ανάπτυξης του συστήματος κομβικό ρόλο έπαιξαν η ανάγκη για γρήγορη απόκριση, η ανάγκη για ασφάλεια και ευκολία κατά την αξιολόγηση αλλά και η προοπτική για διαλειτουργικότητα και επεκτασιμότητα του συστήματος.

Στον πυρήνα της αρχιτεκτονικής βρίσκεται το περιβάλλον Node.js⁹ λόγω της ασύγχρονης φύσης του αλλά και γιατί είναι πολύ δημοφιλές σε τέτοιου είδους εφαρμογές (Tilkon & Vinoski, 2010). Στο περιβάλλον Node.js ο διακομιστής επιτρέπεται να διαχειρίζεται ταυτόχρονα πολλαπλές εξωτερικές κλήσεις προς τα API των γλωσσικών μοντέλων και να μεταφορτώνει εκτενή σώματα κειμένων, χωρίς να επηρεάζεται η απόκριση του συστήματος προς τον χρήστη. Το άρθρωμα Express¹⁰

⁹ <https://nodejs.org/en>

¹⁰ <https://expressjs.com/>

συμπληρώνει το backend και λειτουργεί ως ένας ασφαλής διαμεσολαβητής (server-side proxy).

Το frontend έγινε με χρήση JavaScript¹¹ και αρθρωτή δομή. Αξίζει να σημειωθεί ότι δεν χρησιμοποιήθηκε κάποιο έτοιμο πλαίσιο εργασίας όπως το React¹², ή το Angular¹³. Αν και η χρήση ενός πλαισίου εργασίας όπως τα προηγούμενα θα μπορούσε να επιταχύνει την αρχική ανάπτυξη της διεπαφής του χρήστη (βλέπε Παράρτημα IV), στόχος ήταν η αποφυγή υπολογιστικής επιβάρυνσης και παράλληλα η διασφάλιση της διαφάνειας, της δυνατότητα συντήρησης, διόρθωσης και επεκτασιμότητας του κώδικα. Επίσης, για το σχεδιασμό της διεπαφής επιλέχθηκε το Tailwind CSS¹⁴, ένα πλαίσιο εργασίας που εξασφαλίζει πλήρη προσαρμοστικότητα σε διαφορετικές συσκευές όπως σταθεροί Η/Υ, ταμπλέτες και έξυπνα κινητά, ανεξάρτητα από το λειτουργικό τους σύστημα. Η επιλογή αυτή υποστηρίζει την αρχή της «διάχυτης μάθησης» (ubiquitous learning) (Hwang, 2014) και διασφαλίζει ότι η πρόσβαση στην εφαρμογή είναι ανεξάρτητη από τον τύπο της συσκευής, κάτι σημαντικό για μια εφαρμογή εκμάθησης ξένης γλώσσας.

Τέλος, το backend της εφαρμογής φιλοξενείται στην πλατφόρμα Render¹⁵. Το Render, έπειτα από σύγκριση με άλλες εναλλακτικές λύσεις, όπως το Vercel¹⁶ ή το Netlify¹⁷, προσφέρει ένα πιο απλό, φιλικό και σταθερό περιβάλλον για τη συνεχή λειτουργία εφαρμογών Node.js, χωρίς τους περιορισμούς στους πόρους που θέτουν συχνά οι άλλες πλατφόρμες στην δωρεάν έκδοση. Επιπλέον η πλατφόρμα Render διευκολύνει σημαντικά τη διαχείριση των μεταβλητών περιβάλλοντος, επιτρέποντας την ασφαλή αποθήκευση των API keys για τα γλωσσικά μοντέλα και τα google sheets έξω από τον πηγαίο κώδικα. Η πρακτική αυτή, εκτός από την ασφάλεια που παρέχει, συμμορφώνεται και με τις βασικές αρχές της ερευνητικής δεοντολογίας και προστασίας δεδομένων.

¹¹ <https://vanilla.js.org/>

¹² <https://react.dev/>

¹³ <https://angular.dev/>

¹⁴ <https://tailwindcss.com/>

¹⁵ <https://render.com/>

¹⁶ <https://vercel.com/>

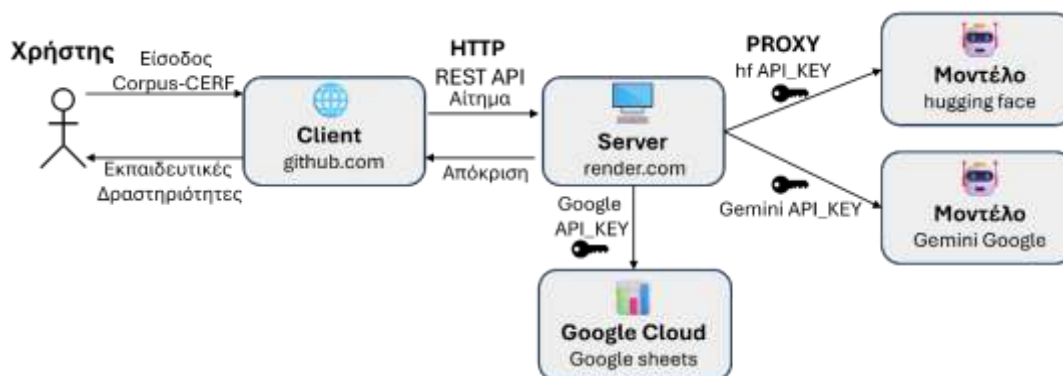
¹⁷ <https://www.netlify.com/>

3.2. Αρχιτεκτονική Υψηλού Επιπέδου

3.2.1. Το μοντέλο Πελάτη-Εξυπηρετητή

Το μοντέλο Πελάτη-Εξυπηρετητή (Client-Server), όπως φαίνεται στην Εικόνα 4, διαχωρίζει σαφώς τη διεπαφή του χρήστη από τη λογική της επεξεργασίας και την επικοινωνία με τα γλωσσικά μοντέλα. Η δομή οργανώνεται γύρω από έναν κεντρικό διακομιστή που λειτουργεί ως διαμεσολαβητής μεταξύ του χρήστη και των εξωτερικών υπηρεσιών.

Στην Εικόνα 4 φαίνεται ότι η αρχιτεκτονική του συστήματος διαρθρώνεται σε τρία διακριτά επίπεδα από αριστερά προς τα δεξιά. Το πρώτο, το επίπεδο του πελάτη, φιλοξενείται στο GitHub¹⁸ σαν ιστοσελίδα και αποτελεί το σημείο αλληλεπίδρασης με τον χρήστη. Σε αυτό το επίπεδο, ορίζεται το παιδαγωγικό πλαίσιο μέσω της εισαγωγής του επιπέδου CEFR και δίνεται το προς ανάλυση λογοτεχνικό κείμενο. Το δεύτερο, το επίπεδο του διακομιστή, το οποίο φιλοξενείται στην πλατφόρμα Render, λειτουργεί ως ο εγκέφαλος του συστήματος. Ο διακομιστής είναι υπεύθυνος για τη ροή, τον έλεγχο και την ασφάλεια των δεδομένων. Τέλος, υπάρχει και το επίπεδο εξωτερικών υπηρεσιών, το οποίο περιλαμβάνει τα Μεγάλα Γλωσσικά Μοντέλα της Google¹⁹ και της Hugging Face²⁰, καθώς και την υποδομή ερευνητικής καταγραφής στο Google Cloud²¹.



Εικόνα 4: Βασική αρχιτεκτονική συστήματος που βασίζεται στο μοντέλο Client-Server. Από αριστερά προς τα δεξιά τα τρία επίπεδα στα οποία χωρίζεται το σύστημα.

¹⁸ <https://github.com/>

¹⁹ <https://aistudio.google.com/>

²⁰ <http://huggingface.co/>

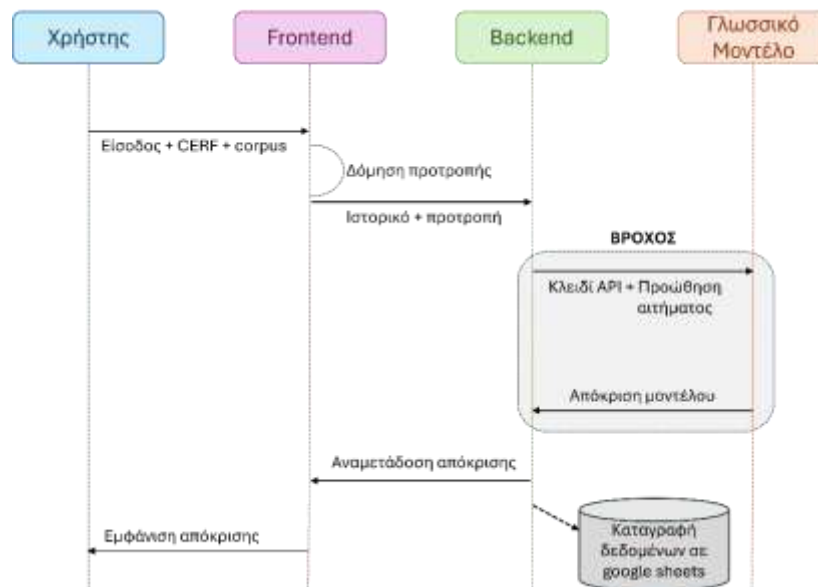
²¹ <https://cloud.google.com/>

Αξίζει να σημειωθεί ότι η λειτουργία του διακομιστή ως proxy εξυπηρετεί διττό σκοπό. Αφενός, διασφαλίζει την ακεραιότητα του συστήματος μέσω της ασφαλούς διαχείρισης των API keys και του ελέγχου των αιτημάτων μακριά από τον πελάτη και αφετέρου, επιτρέπει την αυτοματοποιημένη ερευνητική καταγραφή της μαθησιακής αλληλεπίδρασης σε πραγματικό χρόνο. Έτσι, κάθε αίτημα από τον μαθητή και η αντίστοιχη απόκριση των γλωσσικών μοντέλων αρχειοθετούνται, δημιουργώντας ένα δυναμικό σώμα δεδομένων απαραίτητο για την ποιοτική αξιολόγηση της παιδαγωγικής αποτελεσματικότητας του ψηφιακού βοηθού.

Κλείνοντας, για τις ανάγκες της παρούσας διπλωματικής εργασίας και της πιλοτικής αξιολόγησης, χρησιμοποιήθηκαν αποκλειστικά δωρεάν πακέτα όλων των προαναφερθέντων τεχνολογιών τα οποία όμως και θέτουν συγκεκριμένους περιορισμούς που είναι δίκαιο να αναφέρουμε καθώς η μετάβαση σε μια ευρείας κλίμακας παραγωγική λειτουργία της εφαρμογής, απαιτεί συνδρομητικά πακέτα για την κάλυψη αυξημένων αναγκών σε υπολογιστικούς πόρους και όρια κλήσεων προς τα API.

3.2.2. Ροή Δεδομένων

Η αναλυτική ακολουθία των βημάτων που σταδιακά μετατρέπει την είσοδο του χρήστη σε παιδαγωγικά επεξεργασμένη απόκριση, φαίνεται στο διάγραμμα ακολουθίας της Εικόνας 5. Αρχικά, ο χρήστης εισάγει το αίτημά του μέσω της διεπαφής, επιλέγοντας το επίπεδο CEFR και, προαιρετικά, ένα λογοτεχνικό απόσπασμα. Στη συνέχεια, το frontend συνθέτει το εμπλουτισμένο prompt συνδυάζοντας τις οδηγίες, τον CEFR επίπεδο, το λογοτεχνικό απόσπασμα και το αίτημα του χρήστη και το αποστέλλει μέσω ασφαλούς HTTPS κλήσης στο backend. Εκεί ο διακομιστής επισυνάπτει τα διαπιστευτήρια πρόσβασης και προωθεί το αίτημα στο επιλεγμένο από τον μαθητή γλωσσικό μοντέλο, το οποίο με την σειρά του επεξεργάζεται τα δεδομένα και επιστρέφει την απάντηση. Η απάντηση του γλωσσικού μοντέλου εμφανίζεται στον χρήστη, ενώ παράλληλα καταγράφεται σε Google Sheets. Είναι σημαντικό να αναφέρουμε σε αυτό το σημείο ότι η διαδικασία καταγραφής είναι ασύγχρονη και δεν επηρεάζει την εμπειρία του χρήστη με περιττές καθυστερήσεις (Tilkov & Vinoski, 2010).

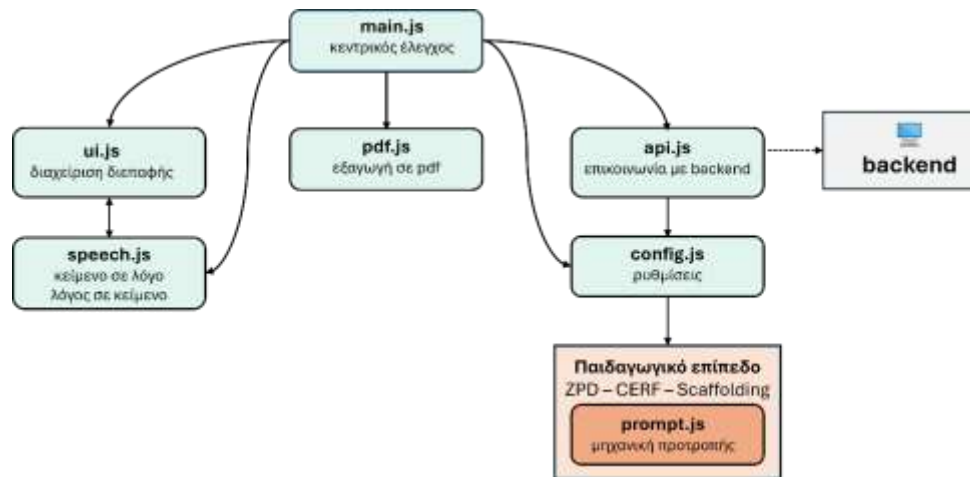


Εικόνα 5: Διάγραμμα ακολουθίας που εξηγεί τον τρόπο αλληλεπίδρασης μεταξύ των οντοτήτων του συστήματος. Στο διάγραμμα φαίνεται η ροή των αιτημάτων από τον χρήστη και των αποκρίσεων του συστήματος καθώς και ο ασύγχρονος μηχανισμός καταγραφής των δεδομένων για την αξιολόγηση.

3.3. Αρχιτεκτονική του Πελάτη (Frontend Modules)

3.3.1. Διαχωρισμός των Αρμοδιοτήτων και Κεντρικός Έλεγχος

Το frontend της εφαρμογής αποτελείται ένα σύνολο από έξι αρθρώματα (javascript αρχεία), που συνεργάζονται μεταξύ τους με σαφώς διαχωρισμένες αρμοδιότητες. Αυτός ο σχεδιασμός έγινε με στόχο την αποφυγή μιας μονολιθικής αρχιτεκτονικής, δηλαδή όλα σε ένα αρχείο, για λόγους εύκολης εκσφαλμάτωσης, συντήρησης αλλά και επεκτασιμότητας του κώδικα. Όπως αποτυπώνεται και στην Εικόνα 6, ο γράφος εξαρτήσεων ακολουθεί ιεραρχική δομή, με μοναδική εξαίρεση την αμφίδρομη σχέση μεταξύ των αρχείων *ui.js* και *speech.js*. Τον ρόλο του κεντρικού ενορχηστρωτή αναλαμβάνει το άρθρωμα *main.js*, το οποίο εισάγει και συντονίζει τις λειτουργίες των επιμέρους μονάδων και αποτελεί το βασικό σημείο ελέγχου της εφαρμογής.



Εικόνα 6: Διάγραμμα αρθρωτής αρχιτεκτονικής (Modular Architecture) του frontend. Στο διάγραμμα απεικονίζεται ο κεντρικός ρόλος του main.js ως ενωρχηστρωτή και ο ιεραρχικός γράφος εξαρτήσεων μεταξύ των επιμέρους λειτουργικών μονάδων της εφαρμογής.

Το άρθρωμα main.js χειρίζεται πληροφορίες όπως, το επιλεγμένο επίπεδο CEFR, το φορτωμένο λογοτεχνικό απόσπασμα, η τελευταία απόκριση του μοντέλου και το ιστορικό της συνομιλίας. Με αυτή την συνειδητή επιλογή, αποφεύγεται η διάχυση της πληροφορίας σε πολλαπλά αρχεία και διασφαλίζεται ότι κάθε αλλαγή στην εφαρμογή ελέγχεται από ένα μόνο κεντρικό σημείο.

Το άρθρωμα ui.js είναι υπεύθυνο για την αλληλεπίδραση με το Document Object Model (DOM). Στο αρχείο ui.js υλοποιήθηκε η συνάρτηση *displayMessage()*, η οποία αποτελεί το σημείο εμφάνισης μηνυμάτων προς τον χρήστη. Κάθε απόκριση του γλωσσικού μοντέλου συνοδεύεται από δυνατότητα ακουστικής αναπαραγωγής, δηλαδή η διεπαφή υποστηρίζεται από φωνητικές λειτουργίες. Παρότι η συνάρτηση αυτή θα μπορούσε να επεκταθεί για την προβολή διαδραστικών στοιχείων (όπως καρτέλες ερωτήσεων), επιλέχθηκε η διατήρηση μιας καθαρής ροής κειμένου. Αυτό έγινε αφενός για να αποφευχθεί η πολυπλοκότητα στη φωνητική αναπαραγωγή λόγω των πολυάριθμων ετικετών HTML και αφετέρου διότι σε υψηλότερα επίπεδα γλωσσομάθειας (C1 και C2), όπου η έμφαση δίνεται στη γραπτή παραγωγή και την ανάλυση, η ανάγκη για τέτοιου είδους οπτικές αναπαραστάσεις είναι περιορισμένη. Παράλληλα, το σύστημα εφαρμόζει μηχανισμό graceful degradation σε περιπτώσεις όπου ο φυλλομετρητής δεν υποστηρίζει αναγνώριση φωνής, η σχετική λειτουργικότητα απενεργοποιείται χωρίς να επηρεάζεται η βασική εμπειρία χρήσης.

3.3.2. Πολυτροπικότητα του Συστήματος

Το αρχείο `speech.js` υλοποιεί τις δύο συμπληρωματικές φωνητικές λειτουργίες. Από τη μία πλευρά, η σύνθεση ομιλίας (Text-to-Speech) χρησιμοποιεί το `SpeechSynthesisUtterance` API με γλώσσα `es-ES`, επιτρέποντας στον μαθητή να ακούσει την απόκριση του μοντέλου με σωστή ισπανική προφορά. Πριν από την ομιλία, εφαρμόζεται καθαρισμός των ειδικών `markdown` χαρακτήρων. Από την άλλη πλευρά, η αναγνώριση ομιλίας (Speech-to-Text) επιτρέπει στον μαθητή να απαντά προφορικά στις ασκήσεις, ενισχύοντας την εξάσκηση της προφορικής παραγωγής. Μεταξύ των μονάδων `ui.js` και `speech.js` υπάρχει αμφίδρομη εξάρτηση, καθώς το πρώτο αξιοποιεί λειτουργίες φωνητικής αναπαραγωγής, ενώ το δεύτερο ενημερώνει δυναμικά τη διεπαφή του χρήστη.

3.3.3. Παραγωγή Εκπαιδευτικού Υλικού

Το αρχείο `pdf.js` αναλαμβάνει να μετατρέψει την ψηφιακή συνομιλία σε ένα στατικό, εκτυπώσιμο έγγραφο. Η λειτουργία αυτή βασίζεται στη βιβλιοθήκη `jsPDF` και φροντίζει για την αυτόματη σελιδοποίηση, την αφαίρεση των εικονιδίων (emoji) και την ενσωμάτωση του λογοτύπου. Το τελικό αρχείο παράγεται σε μορφή A4, έτοιμο για χρήση στη σχολική τάξη ή/και τον διαδραστικό πίνακα. Αυτή η δυνατότητα, αν και απλή, εξυπηρετεί έναν σημαντικό παιδαγωγικό σκοπό γιατί επιτρέπει στο ψηφιακό εργαλείο να «αποσυρθεί» σταδιακά, δίνοντας στον μαθητή ένα φυσικό εκτυπώσιμο έντυπο για αυτόνομη μελέτη. Για την λειτουργία αυτή, υπάρχει ρητή εντολή στο σύστημα προτροπών όπως θα δούμε στην συνέχεια, ώστε να ενεργοποιείται αυτόματα έπειτα από αίτημα του χρήστη, για τη δημιουργία ενός ολοκληρωμένου φύλλου εργασίας.

3.4. Μηχανική Προτροπών (Prompt Engineering)

3.4.1. Καθοδήγηση έναντι Εκπαίδευσης (Prompt vs Fine Tuning)

Ένα από τα κρισιμότερα στάδια αλλά και μια από τις σημαντικότερες προκλήσεις σε τέτοιου είδους εξειδικευμένες εφαρμογές αποτελεί η προσαρμογή του γλωσσικού μοντέλου σε συγκεκριμένα δεδομένα που σχετίζονται με εξειδικευμένα προβλήματα, όπως η παραγωγή ασκήσεων μέσα από λογοτεχνικά αποσπάσματα στην περίπτωσή μας. Η διαδικασία που συνήθως ακολουθείται, είναι γνωστή με τον όρο `fine tuning`, δηλαδή σε ένα ήδη εκπαιδευμένο γλωσσικό μοντέλο ακολουθεί συμπληρωματική

εκπαίδευση χρησιμοποιώντας ένα εξειδικευμένο σύνολο δεδομένων με σκοπό την αναπροσαρμογή των εσωτερικών παραμέτρων (βάρη του νευρωνικού δικτύου) στις ειδικές απαιτήσεις του προβλήματος που θέλουμε να λύσει (Radford et al., 2018). Στην περίπτωση μας, η τεχνική fine tuning απορρίφθηκε για λόγους που σχετίζονται α) με την έλλειψη εξειδικευμένων δεδομένων για την εκπαίδευση και β) με το υπολογιστικό κόστος και τους πόρους που χρειάζεται ένα σύστημα για την διαδικασία της εκπαίδευσης.

Από την άλλη μεριά, η επιλογή της Μηχανικής Προτροπών (Prompt Engineering) βασισμένη σε παραδείγματα (P. Liu et al., 2023), επιτρέπει την επίτευξη του ίδιου παιδαγωγικού στόχου, χωρίς υπολογιστικό κόστος εκπαίδευσης, μετατοπίζοντας την πολυπλοκότητα από το επίπεδο της αρχιτεκτονικής του μοντέλου στο επίπεδο του συντακτικού σχεδιασμού των οδηγιών. Επί της ουσίας, οι οδηγίες προς το γλωσσικό μοντέλο είναι αυτές που του υποδεικνύουν πώς να μεταφράσει την παιδαγωγική θεωρία στην οποία βασίζεται η εφαρμογή, σε ένα λειτουργικό σύστημα εκμάθησης της ισπανικής γλώσσας. Γι' αυτό το λόγο, σχεδιαστικά, οι οδηγίες στην εφαρμογή αποτελούν μια ανεξάρτητη –παραμετροποιήσιμη– μονάδα (prompt.js), που βρίσκεται στο επίπεδο του frontend. Όπως φαίνεται στην Εικόνα 5 και στην Εικόνα 6, από το frontend οι οδηγίες μεταφέρονται στο backend και λειτουργούν ως συνδετικός κρίκος μεταξύ παιδαγωγικής θεωρίας και τεχνολογικής υλοποίησης, επιτρέποντας την ανεξάρτητη εξέλιξη της παιδαγωγικής λογικής, χωρίς αλλαγές στην υπόλοιπη αρχιτεκτονική.

3.4.2. Δομή του Συστήματος Προτροπών

Όσον αφορά τη δομή και τον τρόπο λειτουργίας του συστήματος προτροπών, αυτό βασίζεται στη λογική των λεγόμενων prefix prompts (P. Liu et al., 2023). Δηλαδή, οι οδηγίες και η είσοδος του χρήστη αφού πρώτα συνδυαστούν, σε επόμενο στάδιο διαβιβάζονται στο μοντέλο. Αυτό με την σειρά του επεξεργάζεται το ενιαίο σύνολο και παράγει την τελική απόκριση. Η προσέγγιση αυτή ανήκει στην κατηγορία του tuning-free prompting, καθώς δεν απαιτείται καμία αλλαγή στις εσωτερικές παραμέτρους του μοντέλου.

Φορμαλιστικά, η παραπάνω διαδικασία θα μπορούσε να μοντελοποιηθεί με μια συνάρτηση f_{prompt} που μετατρέπει την αρχική είσοδο x σε μια προτροπή x' , εμπλουτισμένη τις παραμέτρους που σχετίζονται με το αίτημα του μαθητή, το επίπεδο

του μαθητή και το λογοτεχνικό απόσπασμα. Το γλωσσικό μοντέλο επεξεργάζεται την προτροπή x' , με σκοπό να παράξει την απόκριση z του συστήματος. Δηλαδή:

$$z = \text{LLM}(x') = \text{LLM}(f_{\text{prompt}}(x)) = \text{LLM}(f_{\text{prompt}}((x_1, x_2, x_3))), \text{ όπου:}$$

$x = (x_1, x_2, x_3)$:

x_1 : το αίτημα του μαθητή

x_2 : το επίπεδο CEFR

x_3 : το λογοτεχνικό απόσπασμα (προαιρετικό)

$f_{\text{prompt}}(x)$: η συνάρτηση που ενσωματώνει το x στο σύστημα οδηγιών και παράγει x'

LLM: το γλωσσικό μοντέλο

z : η απόκριση που παράγει το μοντέλο.

Σε επίπεδο υλοποίησης, η συνάρτηση f_{prompt} υλοποιείται μέσω της μεθόδου `buildAugmentedPrompt()` που βρίσκεται στο αρχείο `main.js`. Η επιλογή αυτή στον σχεδιασμό, εξασφαλίζει ότι η εφαρμογή παραμένει ανεξάρτητη από το μοντέλο το οποίο καλεί, διατηρώντας την παιδαγωγική λογική, αφού όπως αναφέρθηκε η πολυπλοκότητα δεν εντοπίζεται στα βάρη του νευρωνικού δικτύου, αλλά στην δομή, την σύνταξη και τον τρόπο λειτουργίας του συστήματος προτροπών.

Από την άλλη μεριά, αυτή η δυναμική δόμηση επιτρέπει στο ίδιο σύστημα να παράγει θεμελιωδώς διαφορετικές εκπαιδευτικές εμπειρίες, για διαφορετικά παιδαγωγικά προφίλ χωρίς αλλαγή στον κώδικα, αλλάζοντας μόνο κάποιες από τις παραμέτρους.

3.4.3. Επιλογή Γλώσσας Προτροπής

Η επόμενη σημαντική πρόκληση είναι η γλώσσα στην οποία πρέπει να διατυπωθεί το σύστημα προτροπών. Εφόσον η εφαρμογή απευθύνεται σε ελληνόφωνους μαθητές ισπανικής γλώσσας, φαίνεται εύλογο οι οδηγίες για το μοντέλο να διατυπωθούν στα ελληνικά. Ωστόσο, η σχετική βιβλιογραφία υποδεικνύει μια πιο διαφορετική προσέγγιση.

Οι Huang et al. (2023) απέδειξαν ότι τα γλωσσικά μοντέλα δεν επεξεργάζονται εξίσου αποτελεσματικά όλες τις γλώσσες. Εκπαιδευμένα κατά κύριο λόγο σε αγγλικά γλωσσικά δεδομένα, λόγω της υπεροχής τους στο διαδίκτυο, αποδίδουν καλύτερα όταν οι παιδαγωγικές οδηγίες και η συλλογιστική διατυπώνονται στα αγγλικά. Παράλληλα, οι Shi et al. (2022) κατέδειξαν ότι τα ενδιάμεσα βήματα «σκέψης» των γλωσσικών

μοντέλων είναι αποδοτικότερα στα αγγλικά, από τα αντίστοιχα σε άλλη γλώσσα, λόγω χαμηλής εκπροσώπησης στα δεδομένα εκπαίδευσης.

Βασισμένοι σε αυτά τα ευρήματα, το σύστημα χρησιμοποιεί τρεις γλώσσες: αγγλικά για τη λογική και τις παιδαγωγικές οδηγίες, ελληνικά, με ρητή εντολή, για την έξοδο προς τον χρήστη, ειδικά στα χαμηλότερα CEFR επίπεδα (A1-A2), και ισπανικά αποκλειστικά για ασκήσεις, αποσπάσματα και λογοτεχνικούς όρους. Έτσι, κάθε γλώσσα χρησιμοποιείται με σκοπό το σύστημα να επιτυγχάνει τα βέλτιστα αποτελέσματα στην εκπαιδευτική διαδικασία για την οποία σχεδιάστηκε.

3.4.4. Παιδαγωγική Λογική και Σωκρατικός Διάλογος

Το σημαντικότερο ζήτημα στον σχεδιασμό του συστήματος προτροπών, χωρίς αυτό να αναιρεί την βαρύτητα των προηγούμενων προκλήσεων που συζητήθηκαν στην ενότητα, δεν είναι τεχνικό αλλά αμιγώς παιδαγωγικό. Η απάντηση στο ερώτημα πώς η μηχανική προτροπών κωδικοποιεί όλη την παιδαγωγική θεωρία του 2^{ου} Κεφαλαίου της εργασίας σε λειτουργικές οδηγίες για το μοντέλο. Αυτή, συνοψίζεται σε τέσσερις θεμελιώδεις αρχές που διέπουν το σύστημα:

- i. Μία δραστηριότητα τη φορά. Το μοντέλο παράγει πάντα μία και μόνο δραστηριότητα και περιμένει την απόκριση του μαθητή πριν προχωρήσει στην επόμενη. Με αυτό τον τρόπο προσεγγίζεται η «αρχή του Σωκρατικού διαλόγου», δηλαδή η γνώση δεν μεταδίδεται αλλά ανακαλύπτεται (Paul & Elder, 2008).
- ii. Διαφοροποιημένη ανατροφοδότηση. Το σύστημα δεν αντιμετωπίζει όλες τις απαντήσεις με τον ίδιο τρόπο. Σε σωστή απάντηση ακολουθεί θετική ενίσχυση και αύξηση δυσκολίας για την επόμενη δραστηριότητα. Σε μερικώς σωστή απάντηση αναδεικνύονται τα ορθά στοιχεία και παρέχεται καθοδήγηση. Τέλος, σε περίπτωση λανθασμένης απάντησης, το μοντέλο απαγορεύεται ρητά στο σύστημα οδηγιών να δώσει την ορθή. Αντί αυτού παρέχει μια υπόδειξη, παραπέμπει σε συγκεκριμένο απόσπασμα του κειμένου που κρύβεται η απάντηση ή απλοποιεί επεξηγηματικά την ερώτηση. Πίσω από αυτή την λογική κρύβεται ο μηχανισμός σκαλωσιάς όπως αναλύθηκε στο προηγούμενο κεφάλαιο.
- iii. Σταδιακή απόσυρση (fading). Καθώς ο μαθητής αποκτά μεγαλύτερη εξοικείωση, το σύστημα έχει οδηγία να μειώνει προοδευτικά την καθοδήγηση και να εισάγει ερωτήσεις ανοικτού τύπου που επιτρέπουν πολλαπλές ερμηνείες. Αυτή η δυναμική

μετάβαση προς την αυτονομία του μαθητή υλοποιεί την αρχή της ZEA (Cole & Scribner, 1978) στην πράξη.

- iv. Μαζική παραγωγή υλικού. Όταν ο χρήστης ζητά ρητά ένα φύλλο εργασίας για εξαγωγή σε μορφότυπο pdf, το σύστημα αναστέλλει τον διαλογικό τρόπο λειτουργίας του και παράγει ένα δομημένο σύνολο από δραστηριότητες κατάλληλες για εκτύπωση (Wei et al., 2022; White et al., 2023). Η λογική πίσω από αυτή την εξαίρεση, συνδέεται άμεσα με την αρχή της απόσυρσης, καθώς ο μαθητής λαμβάνει υλικό για αυτόνομη μελέτη, εκτός συστήματος και σε δεύτερο χρόνο. Το πρόβλημα που παρουσιάζεται σε αυτή την περίπτωση είναι η αδυναμία ανατροφοδότησης του συστήματος προς τον μαθητή, αφού το υλικό μπορεί να μην να φορτωθεί στο σύστημα, αλλά έχει χαθεί όλο το ιστορικό συνομιλίας (λογοτεχνικό απόσπασμα, αίτημα μαθητή και επίπεδο CEFR).

Η συνολική αυτή προσέγγιση ευθυγραμμίζεται σε απόλυτο βαθμό με τα ερευνητικά δεδομένα που καταδεικνύουν ότι η μετάβαση από το παράδειγμα «Ερώτηση-Απάντηση» σε ένα «Thought-Provoking» μοντέλο διδασκαλίας ενισχύει την ουσιαστική κατανόηση και την ανάπτυξη δεξιοτήτων από την πλευρά του μαθητή, όπως η ικανότητα επίλυσης προβλημάτων (J. Liu et al., 2024).

3.4.5. Ιστορικό Συνομιλίας και το Prompt ως αντικείμενο έρευνας

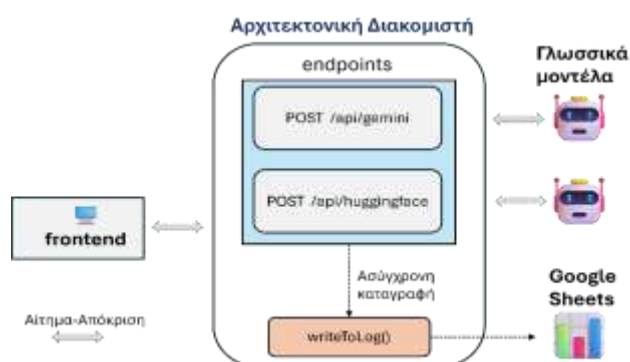
Η διατήρηση του ιστορικού των συνομιλιών αποτελεί τεχνική πρόκληση, λόγω της ασύμβατης δομής των δύο γλωσσικών μοντέλων που χρησιμοποιούνται στο σύστημα. Το Gemini χρησιμοποιεί nested array format, ενώ το HuggingFace/OpenAI-style API χρησιμοποιεί flat string format. Για τον λόγο αυτό διατηρούνται δύο ανεξάρτητα ιστορικά — ένα ανά μοντέλο — που αρχικοποιούνται κατά την πρώτη αλληλεπίδραση και σταδιακά εμπλουτίζονται σε κάθε ανταλλαγή. Η επιλογή αυτή έχει και παιδαγωγική διάσταση διότι η διατήρηση του ιστορικού μιμείται τη συνεχή αλληλεπίδραση μαθητή-δασκάλου.

Τέλος, το σύστημα προτροπών, εκτός από παιδαγωγικό, αποτελεί ταυτόχρονα ερευνητικό εργαλείο. Η εφαρμογή καταγράφει κάθε αλληλεπίδραση στο Google Sheets, δημιουργώντας ένα σώμα δεδομένων που επιτρέπει την μελλοντική της αξιολόγηση. Δηλαδή πώς διαφορετικές διατυπώσεις επηρεάζουν την ποιότητα της Σωκρατικής αλληλεπίδρασης, πού εντοπίζονται αστοχίες του συστήματος στον παιδαγωγικό διάλογο, ποια επίπεδα CEFR παράγουν τις πιο ενδιαφέρουσες ακολουθίες

συζητήσεων κ.λπ. Η αρχιτεκτονική δηλαδή, επιτρέπει η μηχανική προτροπών να εξελιχθεί από απλή τεχνική επιλογή σε αντικείμενο επιστημονικής διερεύνησης (Holmes et al., 2019).

3.5. Η αρχιτεκτονική του Διακομιστή

Η Εικόνα 7 επεκτείνει το διάγραμμα της Εικόνα 4 και παρουσιάζει την αρχιτεκτονική του διακομιστή, ο πυρήνας του οποίου αποτελείται από τα 2 επιμέρους endpoints και τον μηχανισμό διαχείρισης και καταγραφής αιτημάτων. Η επιλογή αυτή της αρχιτεκτονικής εξυπηρετεί τη δυνατότητα συγκριτικής αξιολόγησης των γλωσσικών μοντέλων και την επεκτασιμότητα της εφαρμογής ως προς την ενσωμάτωση νέων μοντέλων στο μέλλον. Η προσθήκη ενός νέου μοντέλου στο σύστημα, απαιτεί απλώς τη δημιουργία ενός νέου endpoint, χωρίς αλλαγές στο frontend της εφαρμογής ή στο σύστημα οδηγιών.



Εικόνα 7: Εσωτερική αρχιτεκτονική διακομιστή. Στο εσωτερικό του φαίνονται τα δύο endpoints επικοινωνίας με τα γλωσσικά μοντέλα και ο μηχανισμός ασύγχρονης καταγραφής `writeToLog()` των δεδομένων στο Google Sheets για ερευνητική αξιολόγηση.

Ως προς την ασφάλεια, κανένα διαπιστευτήριο δεν εκτίθεται στον φυλλομετρητή, αφού τα κλειδιά API αποθηκεύονται αποκλειστικά ως μεταβλητές περιβάλλοντος στον διακομιστή (Render), έξω από τον πηγαίο κώδικα, κάτι που ευθυγραμμίζεται και με τις αρχές ερευνητικής δεοντολογίας. Ο μηχανισμός καταγραφής υλοποιείται μέσω της μεθόδου `writeToLog()`, η οποία αρχειοθετεί αυτόματα κάθε αλληλεπίδραση στο Google Sheets αποθηκεύοντας την χρονική σφραγίδα, το μοντέλο, το αίτημα του μαθητή και την απόκριση. Το όριο μεταφοράς δεδομένων στον διακομιστή, για την υποστήριξη της ανάλυσης των λογοτεχνικών αποσπασμάτων, ορίστηκε στα 50MB, διατηρώντας την δυνατότητα επεκτασιμότητας της εφαρμογής σε πιο σύνθετα ερευνητικά σενάρια.

Κλείνοντας, αξίζει να αναφερθούν οι βασικοί περιορισμοί αυτής της αρχιτεκτονικής επιλογής. Η εφαρμογή εξαρτάται από εξωτερικά APIs, τα οποία έχουν τους δικούς τους όρους χρήσης και περιορισμούς. Επιπλέον, η πιθανότητα παραγωγής εσφαλμένων ή ανακριβών αποκρίσεων από τα μοντέλα, καθιστά απαραίτητη την ανθρώπινη επίβλεψη. Τέλος, η απουσία μόνιμης μνήμης μεταξύ των συνεδριών περιορίζει τη διαχρονική παρακολούθηση της προόδου του μαθητή.

Κεφάλαιο 4: Μεθοδολογία Έρευνας και Αξιολόγησης

4.1. Μεθοδολογία Αξιολόγησης

4.1.1. Μεθοδολογικό Πλαίσιο

Σκοπός της αξιολόγησης είναι να διερευνηθεί κατά πόσον το προτεινόμενο σύστημα ΤΝ επιτυγχάνει τους παιδαγωγικούς του στόχους, δηλαδή την υποστήριξη της μαθησιακής διαδικασίας μέσω της σκαλωσιάς και του Σωκρατικού διαλόγου στην διδασκαλία της ισπανικής γλώσσας μέσα από λογοτεχνικά κείμενα. Επίσης, κατά την αξιολόγηση εκτιμάται η χρηστικότητα και ο βαθμός αποδοχής της εφαρμογής από μαθητές και εκπαιδευτικούς. Η διαδικασία εντάσσεται στο πλαίσιο *Computer-Assisted Language Learning* (CALL), στο οποίο η τεκμηριωμένη αξιολόγηση τεχνολογικών εργαλείων αποτελεί κεντρική μεθοδολογική απαίτηση (Chapelle & Chapelle, 2001).

Η αξιολόγηση περιστρέφεται γύρω από τρία ερευνητικά ερωτήματα τα οποία αντιστοιχούν στη γνωστική αποτελεσματικότητα, την αποδοχή των χρηστών και τον βαθμό στον οποίο το σύστημα υλοποιεί στην πράξη τις παιδαγωγικές αρχές που το διέπουν, όπως ορίστηκαν στα κεφάλαιο 2. Η διαδικασία ακολουθεί την στρατηγική τριγωνισμού των δεδομένων με βασικό στόχο την εξασφάλιση της εγκυρότητας των αποτελεσμάτων.

4.1.2. Ερευνητικά Ερωτήματα

Τα ερευνητικά ερωτήματα που επιχειρείται να απαντηθούν είναι τα ακόλουθα:

EE1: Σε ποιο βαθμό η αλληλεπίδραση με το σύστημα συμβάλλει, κατά την αντίληψη των μαθητών, στην υποστήριξη της κατανόησης του λογοτεχνικού κειμένου και στην ανάπτυξη γλωσσικών δεξιοτήτων ανά επίπεδο;

EE2: Πώς αξιολογούν οι μαθητές και οι εκπαιδευτικοί την παιδαγωγική αποτελεσματικότητα και τη χρηστικότητα του συστήματος;

EE3: Κατά πόσον το σύστημα υλοποιεί στην πράξη τις αρχές της σκαλωσιάς, της σταδιακής απόσυρσης υποστήριξης και του Σωκρατικού διαλόγου, διασφαλίζοντας παράλληλα την παιδαγωγική εγκυρότητα μέσω του περιορισμού φαινομένων ψευδαισθήσεων και αλγοριθμικής μεροληψίας;

4.2. Ερευνητικός Σχεδιασμός

Κατά την αξιολόγηση ενός εκπαιδευτικού ψηφιακού συστήματος, εκτός από τα αριθμητικά μετρήσιμα αποτελέσματα εξίσου σημαντικά κριτήρια αποτελούν η εμπειρία του μαθητή και η άποψη του ειδικού (παιδαγωγού), μέτρα που δύσκολα αποτυπώνονται σε απόλυτη αριθμητική κλίμακα. Γι' αυτό, για την απάντηση στα ερευνητικά ερωτήματα που διατυπώθηκαν, επιλέχθηκε μεικτή μεθοδολογία που συνδυάζει τόσο ποσοτικά όσο και ποιοτικά δεδομένα (Creswell & Clark, 2007).

Συγκεκριμένα, ακολουθείται ο συγκλίνων παράλληλος σχεδιασμός. Αυτό σημαίνει ότι τα ποσοτικά δεδομένα (βαθμολογίες και αριθμητικές κλίμακες) συλλέγονται ταυτόχρονα με τα ποιοτικά δεδομένα (ερωτήσεις ανοικτού τύπου σε καθηγητές και μαθητές). Στη συνέχεια, αυτά τα δεδομένα διασταυρώνονται μέσω της στρατηγικής του τριγωνισμού. Με αυτόν τον τρόπο, ελέγχεται αν τα αριθμητικά αποτελέσματα επιβεβαιώνονται στην πράξη μέσα από την ατομική εμπειρία των χρηστών, αυξάνοντας έτσι την αξιοπιστία της έρευνας (Denzin, 2012).

Η λογική πίσω από αυτή την προσέγγιση είναι κάθε φορά να επιλέξουμε τα κατάλληλα εργαλεία αξιολόγησης που εξυπηρετούν καλύτερα τα ερευνητικά ερωτήματα που ετέθησαν. Έτσι, οι αριθμοί μας δείχνουν «πόσο» αποτελεσματικό είναι το σύστημα και οι ποιοτικές αναλύσεις μέσα από ελεύθερες απαντήσεις σε ανοικτού τύπου ερωτήσεις μας βοηθάνε να καταλάβουμε το «πώς» και το «γιατί» λειτούργησαν – ή δεν λειτούργησαν – οι παιδαγωγικές μας επιλογές στην πράξη.

4.3. Εργαλεία Συλλογής Δεδομένων και Δειγματοληψία

4.3.1. Ερωτηματολόγια

Πριν την έναρξη της διαδικασίας της αξιολόγησης, οι μαθητές συμπλήρωσαν ένα ερωτηματολόγιο δημογραφικού ενδιαφέροντος (βλ. Παράρτημα Ι). Το ερωτηματολόγιο περιέχει ερωτήσεις σχετικά με την ηλικία, το επίπεδο σπουδών, το επίπεδο ισπανικής, τα κίνητρα εκμάθησης, την εμπειρία με άλλες ξένες γλώσσες και την προηγούμενη χρήση ψηφιακών εργαλείων στην εκμάθηση ξένης γλώσσας. Τέλος, περιλαμβάνει μια ερώτηση βάση, σχετικά με την συμβολή της ΤΝ στη γλωσσομάθεια, η οποία λειτουργεί ως σημείο σύγκρισης με τις απαντήσεις των χρηστών μετά τη χρήση του συστήματος.

Μετά την αλληλεπίδραση των χρηστών με το σύστημα, για την ανάλυση των τριών ερευνητικών ερωτημάτων, χρησιμοποιήθηκαν δύο εργαλεία συλλογής ποσοτικών και ποιοτικών δεδομένων, ένα ερωτηματολόγιο που συμπληρώθηκε από τους μαθητές και ένα από τους καθηγητές που επέβλεψαν την διαδικασία.

Το ερωτηματολόγιο που συμπληρώθηκε από τους μαθητές αποτελείται από τρία μέρη (βλ. Παράρτημα II). Το πρώτο μέρος περιλαμβάνει οκτώ ερωτήσεις κλειστού τύπου σε κλίμακα Likert 1-5, εμπνευσμένες από το Technology Acceptance Model (TAM; Davis, 1989) και το User Experience Questionnaire Short (UEQ-S; Schrepp et al., 2017), και μετρά τη χρηστικότητα της εφαρμογής, την αντιστοιχία με το CEFR επίπεδο και τη συνολική ικανοποίηση του μαθητή. Το δεύτερο μέρος περιλαμβάνει τέσσερις επιπλέον ερωτήσεις τύπου Likert που εστιάζουν στη μαθησιακή εμπλοκή, δηλαδή την αίσθηση ενεργούς συμμετοχής, την ασφάλεια έκφρασης και την σύνδεση με προηγούμενη γνώση του μαθητή. Το τελευταίο μέρος περιλαμβάνει τέσσερις ερωτήσεις ανοικτού τύπου, που αποτυπώνουν την προσωπική εμπειρία του μαθητή. Δηλαδή, οι ανοιχτές ερωτήσεις εξετάζουν τι άρεσε και τι δεν λειτούργησε στο σύστημα συγκριτικά με τον καθηγητή, καθώς και σχετικές προτάσεις βελτίωσης.

Το δεύτερο εργαλείο συλλογής δεδομένων απευθύνεται στους εκπαιδευτικούς (βλ. Παράρτημα III) και αποτελείται επίσης από τρία μέρη. Το πρώτο μέρος περιλαμβάνει έντεκα ερωτήσεις κλειστού τύπου σε κλίμακα Likert 1-5, οργανωμένες σε τέσσερις άξονες: κατά πόσο εφαρμόστηκε ο Σωκρατικός διάλογος και σκαλωσιά στην πράξη (E1-E4) – ερωτήσεις που αποτελούν την εξειδίκευση των αρχών του Vygotsky όπως αναπτύχθηκαν στο Κεφάλαιο 2 –, ποια είναι η αντιστοιχία με το CEFR επίπεδο (E5-E7), ποια είναι ποιότητα ανατροφοδότησης του συστήματος (E8 και E9) και κατά πόσο το περιεχόμενο που παράγει το σύστημα είναι αξιόπιστο (E10 και E11).

Το δεύτερο μέρος του ερωτηματολογίου των καθηγητών αποτελείται από τρεις ερωτήσεις, μέσω των οποίων οι εκπαιδευτικοί αξιολογούν συνδυαστικά την παιδαγωγική αποτελεσματικότητα του συστήματος, τη γλωσσική του ακρίβεια σε σχέση με το επίπεδο CEFR των μαθητών και τον βαθμό εμπλοκής που προκάλεσε στον μαθητή. Τέλος, το τρίτο μέρος αποτελεί τον οδηγό μιας δομημένης συνέντευξης με τέσσερις ερωτήσεις ανοικτού τύπου που καλύπτουν την παιδαγωγική λογική, τη γλωσσική ακρίβεια, τη μαθησιακή εμπλοκή και το πλαίσιο εφαρμογής του συστήματος στην τάξη.

4.3.2. Σώμα Κειμένων

Η επιλογή του corpus για την αξιολόγηση του συστήματος, πέρα από τεχνική (έκταση του αποσπάσματος, σωστή κωδικοποίηση κειμένου κ.λπ.) είναι και αμιγώς παιδαγωγική απόφαση την οποία καθόρισαν οι εκπαιδευτικοί που συμμετείχαν στην έρευνα. Προϋπόθεση για την δοκιμή του συστήματος σε συνθήκες που έχουν πραγματική παιδαγωγική αξία αποτελούν κείμενα με γλωσσική αντιστοιχία στο CEFR επίπεδο, λεξιλογικό πλούτο, πολιτισμικό βάθος και λογοτεχνική ταυτότητα.

Για τα επίπεδα A1-A2 επιλέχθηκε το ποίημα «*Oda a la papa*» του Pablo Neruda (1954) από τη συλλογή «*Odas Elementales*». Το ποίημα χαρακτηρίζεται από απλότητα — σύντομοι στίχοι, καθημερινά αντικείμενα και λεξιλόγιο της καθημερινότητας —, αμεσότητα και χιουμοριστικό ύφος, το οποίο ο Neruda επέλεξε συνειδητά για να γράψει ποίηση προσβάσιμη στον απλό άνθρωπο (Thakur, 2024). Αυτά τα χαρακτηριστικά ευθυγραμμίζονται με τις απαιτήσεις του επιπέδου A1-A2. Ωστόσο, η πατάτα ως σύμβολο λατινοαμερικάνικης ταυτότητας και αυτόχθονης κουλτούρας προσδίδει ένα ερμηνευτικό βάθος που υποχρεώνει το σύστημα να καθοδηγήσει τον αρχάριο μαθητή να αντιληφθεί τη μεταφορά, χωρίς να χρησιμοποιήσει δύσκολη γλώσσα και χωρίς να του προσφέρει έτοιμες απαντήσεις.

Για τα επίπεδα B1-B2 επιλέχθηκε το διήγημα «*La luz es como el agua*» (García Márquez, 1992) από τη συλλογή «*Doce cuentos peregrinos*». Το απόσπασμα χαρακτηρίζεται από σύντομες προτάσεις, γραμμική πλοκή και καθημερινό λεξιλόγιο που αντιστοιχεί στο επίπεδο B1-B2. Επίσης, η έντονη πολιτισμική του διάσταση, καθώς το κείμενο πραγματεύεται έννοιες όπως η νοσταλγία για την πατρίδα, η λατινοαμερικάνικη ταυτότητα, και η έντονη αντίθεση με την ευρωπαϊκή πραγματικότητα (Bell-Villada, 1990), εξυπηρετεί τον στόχο του συστήματος να συνδέσει το κείμενο με το πολιτισμικό πλαίσιο. Στο απόσπασμα, που ανήκει στο λογοτεχνικό κίνημα του Μαγικού Ρεαλισμού, παρατηρείται μια έντονη μεταφορά – το φως που ρέει σαν νερό – χωρίς καμία επεξήγηση (Faris, 2004). Εδώ η πρόκληση για το σύστημα είναι η μεταφορά να αντιμετωπιστεί ως στοιχείο που προσδίδει πολιτισμική ταυτότητα στο απόσπασμα και όχι ως λογικό σφάλμα, αποφεύγοντας παράλληλα την αλγοριθμική μεροληψία απέναντι στη λατινοαμερικάνικη κουλτούρα.

Για το επίπεδο C1-C2 επιλέχθηκε το διήγημα «*Casa tomada*» (Cortázar, 1951) από τη συλλογή «*Bestiario*». Το κείμενο χαρακτηρίζεται από σκόπιμη αοριστία και από

πολλαπλές ερμηνευτικές δυνατότητες τις οποίες ο ίδιος ο συγγραφέας σκόπιμα αρνείται να ιεραρχήσει. Το κείμενο ανήκει στο λατινοαμερικάνικο φανταστικό διήγημα στο οποίο το ανεξήγητο εισβάλλει στο καθημερινό, δημιουργώντας μια συναισθηματική αμφιταλάντευση και μια διαρκή αγωνία στον αναγνώστη (Standish, 2001). Αυτά τα χαρακτηριστικά του αποσπάσματος, το καθιστούν ιδανικό για τα επίπεδα ανώτερα επίπεδα, όπου ο μαθητής αναμένεται να μπορεί να αξιολογεί διαφορετικές ερμηνευτικές προσεγγίσεις. Ταυτόχρονα, αποτελεί τεράστια πρόκληση για το σύστημα γιατί δοκιμάζεται η δυνατότητά του να παρουσιάσει τις πολλαπλές ερμηνείες, αποφεύγοντας τις ψευδαισθήσεις.

Να σημειωθεί ότι στο σύστημα εισήχθησαν σύντομα αποσπάσματα από τα παραπάνω έργα και όχι ολόκληρα τα κείμενα. Η χρήση αυτή εντάσσεται στο πλαίσιο της εκπαιδευτικής και ερευνητικής χρήσης για μη εμπορικούς σκοπούς.

4.3.3. Δειγματοληψία

Η αξιολόγηση του συστήματος πραγματοποιήθηκε με τη συμμετοχή μαθητών και εκπαιδευτικών από τα κέντρα ξένων γλωσσών «Ελφη Ασλάνη» και «FEDIS» στα Ιωάννινα. Η επιλογή του δείγματος έγινε με τη μέθοδο της δειγματοληψίας ευκολίας, αξιοποιώντας την πρόσβαση σε μαθητές που ήδη διδάσκονταν την ισπανική γλώσσα σε διαφορετικά επίπεδα. Η μέθοδος αυτή επιλέχθηκε λόγω της φύσης της παρούσας έρευνας (πilotική μελέτη), στην οποία η πρόσβαση σε κατάλληλο δείγμα αποτελεί πρωταρχική προϋπόθεση (Cohen et al., 2002).



Εικόνα 8: Στιγμιότυπα από τη διαδικασία αξιολόγησης στα φροντιστήρια, όπου οι μαθητές αλληλεπιδρούν με την εφαρμογή μέσω φορητών συσκευών υπό την καθοδήγηση της εκπαιδευτικού.

Μέρος από την πρακτική εφαρμογή της διαδικασίας αποτυπώνεται στην Εικόνα 8. Συνολικά, στην έρευνα συμμετείχαν 12 μαθητές, οι οποίοι κατανεμήθηκαν σε τρεις

ομάδες με βάση το επίπεδο γλωσσομάθειάς τους (4 μαθητές στα επίπεδα A1-A2, 6 μαθητές στα επίπεδα B1-B2 και 2 μαθητές στα επίπεδα C1-C2). Παράλληλα, την εφαρμογή αξιολόγησαν 2 έμπειροι εκπαιδευτικοί των παραπάνω φροντιστηρίων. Οι εκπαιδευτικοί, που είχαν τον ρόλο του ειδικού παρατηρητή, εστίασαν στην παιδαγωγική ορθότητα των απαντήσεων της TN και στην ανίχνευση τυχόν σφαλμάτων ή μεροληψιών. Επιπλέον, επέβλεψαν την διαδικασία με σκοπό να επέμβουν όπου κρίθηκε απαραίτητο.

Η ερευνητική διαδικασία για την αξιολόγηση του συστήματος υλοποιήθηκε σε τρεις διακριτές φάσεις. Η έρευνα διεξήχθη σύμφωνα με τις αρχές ηθικής δεοντολογίας αφού κατά την προετοιμασία, πριν από την έναρξη της αλληλεπίδρασης, οι μαθητές ενημερώθηκαν για τον σκοπό της μελέτης και τη διασφάλιση της ανωνυμίας τους και παρείχαν ρητά τη συγκατάθεσή τους. Στη συνέχεια, συμπλήρωσαν το ερωτηματολόγιο δημογραφικού προφίλ και πραγματοποιήθηκε μια πολύ σύντομη επίδειξη της διεπαφής του συστήματος για την εξοικείωσή τους με το ψηφιακό περιβάλλον. Έγινε διανομή των κειμένων ανά επίπεδο και ακολούθησε η ανάγνωσή τους από τους μαθητές.

Στη φάση της παρέμβασης, οι μαθητές αλληλεπέδρασαν ατομικά με τον ψηφιακό βοηθό μέσω φορητών συσκευών (laptops και smartphones) όπως φαίνεται στην Εικόνα 8, μελετώντας το λογοτεχνικό απόσπασμα που αντιστοιχούσε στο επίπεδό τους. Κατά τη διάρκεια της συνεδρίας (διάρκειας περίπου 30 λεπτών ανά γλωσσικό επίπεδο), οι εκπαιδευτικοί-αξιολογητές λειτουργούσαν ως κριτικοί παρατηρητές, καταγράφοντας σε πραγματικό χρόνο τις αντιδράσεις των μαθητών και την ποιότητα των αποκρίσεων της TN.

Τέλος, στη φάση της αποτίμησης, μετά την ολοκλήρωση της αλληλεπίδρασης, οι μαθητές συμπλήρωσαν τα αντίστοιχα ερωτηματολόγια αξιολόγησης της εφαρμογής. Οι εκπαιδευτικοί μελέτησαν τα αρχεία καταγραφής με τις αποκρίσεις του συστήματος στις ερωτήσεις των μαθητών και συμπλήρωσαν τα αντίστοιχα ερωτηματολόγια που αφορούν στους εκπαιδευτικούς όπως αυτά παρουσιάστηκαν στην ενότητα 4.3.1.

4.4. Ανάλυση Δεδομένων

4.4.1. Ποσοτική Ανάλυση

Για την ανάλυση των ποσοτικών δεδομένων που συλλέχθηκαν από τα ερωτηματολόγια μαθητών και εκπαιδευτικών υπολογίστηκαν οι βασικές περιγραφικές στατιστικές

(μέσος όρος, τυπική απόκλιση) ανά ερώτηση και ανά υποομάδα (επίπεδο CEFR, γλωσσικό μοντέλο). Οι υπολογισμοί των τιμών έγιναν με script στην γλώσσα προγραμματισμού python σε περιβάλλον jupyter notebook.

Για τον υπολογισμό της συμφωνίας μεταξύ των δύο εκπαιδευτικών-αξιολογητών χρησιμοποιήθηκε το ποσοστό απόλυτης συμφωνίας (Stemler, 2004). Η επιλογή αυτή κρίθηκε κατάλληλη λόγω του μικρού αριθμού αξιολογητών ($N=2$), που δεν επιτρέπει τη χρήση πιο σύνθετων δεικτών συμφωνίας.

4.4.2. Ποιοτική Ανάλυση

Για την ανάλυση των ποιοτικών δεδομένων — ερωτήσεις ανοικτού τύπου μαθητών και εκπαιδευτικών, καθώς και αρχεία καταγραφής των συνεδριών (chat logs) — εφαρμόστηκε θεματική ανάλυση σύμφωνα με τη μεθοδολογία των Braun & Clarke (2006). Πρόκειται για μια ευέλικτη ποιοτική μέθοδο που επιτρέπει την αναγνώριση, ανάλυση και παρουσίαση επαναλαμβανόμενων μοτίβων μέσα στα δεδομένα.

Οι απαντήσεις κωδικοποιήθηκαν επαγωγικά, δηλαδή οι θεματικές κατηγορίες αναδύθηκαν από τα ίδια τα δεδομένα και όχι από προκαθορισμένο θεωρητικό πλαίσιο. Η κωδικοποίηση πραγματοποιήθηκε χειροκίνητα με τη χρήση πινάκων καταγραφής, μέθοδος απόλυτα κατάλληλη για το μέγεθος του δείγματος της παρούσας πιλοτικής μελέτης. Από την ανάλυση προέκυψαν τέσσερις θεματικές κατηγορίες που εμφανίστηκαν συστηματικά τόσο στις απαντήσεις των μαθητών όσο και στις παρατηρήσεις των εκπαιδευτικών και στα αρχεία καταγραφής: α) αξιολόγηση του Σωκρατικού διαλόγου, β) τεχνικά προβλήματα, γ) σύγκριση με τον καθηγητή και δ) προτάσεις βελτίωσης.

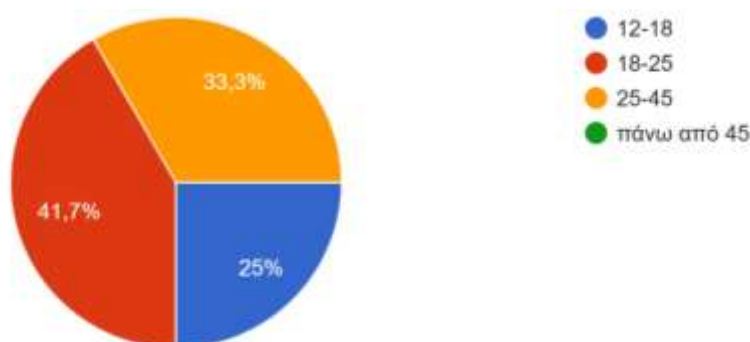
4.4.3. Τριγωνισμός Δεδομένων

Τα ποσοτικά και ποιοτικά δεδομένα δεν αναλύθηκαν ανεξάρτητα αλλά διασταυρώθηκαν μέσω της στρατηγικής του τριγωνισμού (Denzin, 2012). Συγκεκριμένα, τα αποτελέσματα των ερωτηματολογίων (ποσοτικά) αντιπαραβλήθηκαν με τις ανοιχτές απαντήσεις των μαθητών (ποιοτικά), τις παρατηρήσεις των εκπαιδευτικών και τα αρχεία καταγραφής των συνεδριών. Με αυτόν τον τρόπο, κάθε εύρημα επιβεβαιώνεται — ή αμφισβητείται — από τουλάχιστον δύο ανεξάρτητες πηγές δεδομένων, ενισχύοντας την εγκυρότητα των συμπερασμάτων.

Κεφάλαιο 5: Αποτελέσματα Αξιολόγησης

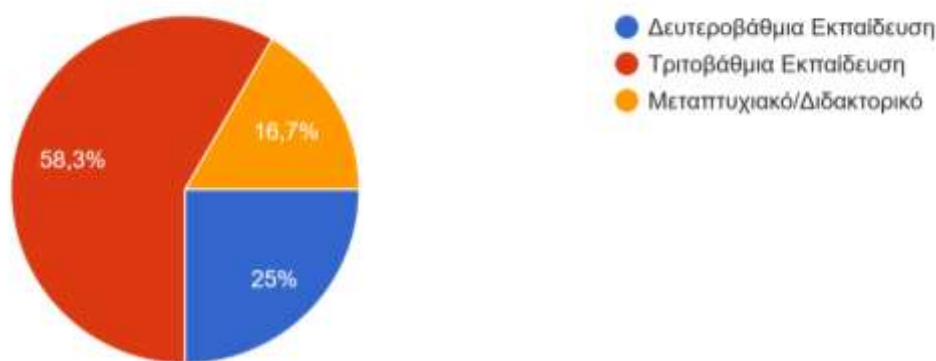
5.1. Προφίλ Χρηστών

Στην έρευνα συμμετείχαν συνολικά δώδεκα μαθητές (M1-M12) ισπανικής γλώσσας. Οι μαθητές, πριν από τη χρήση του συστήματος, συμπλήρωσαν το ερωτηματολόγιο που σχετίζεται με το προφίλ του χρήστη (βλέπε Παράρτημα Ι). Το δείγμα καλύπτει τρεις ηλικιακές ομάδες: πέντε συμμετέχοντες ανήκουν στην ομάδα 18-25 ετών (41,7%), τέσσερις στην ομάδα 25-45 ετών (33,3%) και τρεις στην ομάδα 12-18 ετών (25%) όπως αποτυπώνεται και στην Εικόνα 9.



Εικόνα 9: Ηλικιακή κατανομή χρηστών του συστήματος.

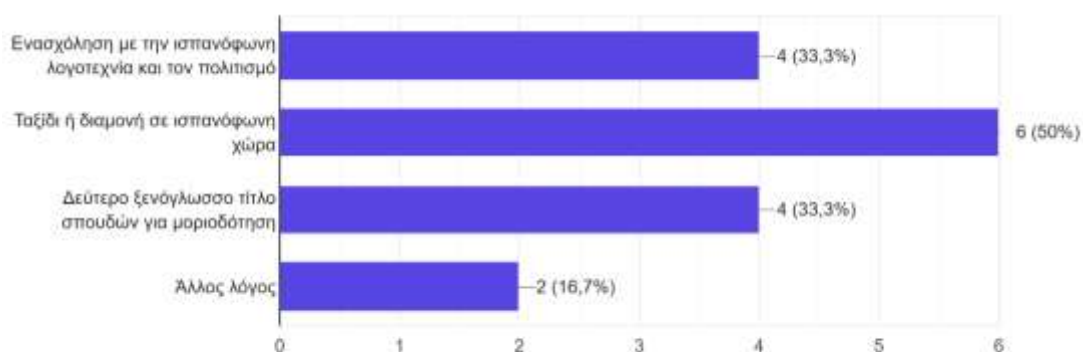
Ως προς το επίπεδο σπουδών, η πλειονότητα των μαθητών προέρχεται από την τριτοβάθμια εκπαίδευση (9 άτομα, 75%), με 2 εξ αυτών να κατέχουν μεταπτυχιακό ή διδακτορικό τίτλο σπουδών, και τρεις από τη δευτεροβάθμια εκπαίδευση (25%) όπως φαίνεται στην Εικόνα 10. Κατά συνέπεια, η ηλικιακή και γνωστική ετερογένεια συνολικά βοήθησε να αξιολογηθεί το σύστημα από διαφορετικά προφίλ.



Εικόνα 10: Επίπεδο σπουδών των χρηστών του συστήματος που συμμετείχαν στην αξιολόγηση.

Ως προς το επίπεδο της ισπανικής γλώσσας των χρηστών, η αξιολόγηση σχεδιάστηκε ώστε να καλύπτει και τα τρία επίπεδα του CEFR αφού, έξι συμμετέχοντες (M7-M12) δήλωσαν ότι φοιτούν στο επίπεδο B1-B2 (50%), τέσσερις (M1-M4) στο A1-A2 (33,3%) και δύο (M5 και M6) στο C1-C2 (16,7%).

Όσον αφορά στα κίνητρα εκμάθησης, άξιο αναφοράς είναι ότι η πολιτισμική διάσταση, που αποτελεί έναν από τους κεντρικούς άξονες του σχεδιασμού του συστήματος, είναι ανάμεσα στα βασικά κίνητρα όπως φαίνεται και στο γράφημα που απεικονίζεται στην Εικόνα 11. Η πιο συχνή απάντηση που έδωσαν οι μαθητές στην ερώτηση γιατί επέλεξαν την εκμάθηση της ισπανικής ήταν το ταξίδι ή η διαμονή σε ισπανόφωνη χώρα (6 αναφορές), ακολουθούμενη από την ενασχόληση με την ισπανόφωνη λογοτεχνία και τον πολιτισμό (4 αναφορές). Τέλος, τέσσερις μαθητές απάντησαν ότι η εκμάθηση μιας ξένης γλώσσας τους ενδιαφέρει για ενδεχόμενη μοριοδότηση σε διαγωνισμούς και δύο μαθητές ανέφεραν άλλο λόγο χωρίς περαιτέρω εξειδίκευση.

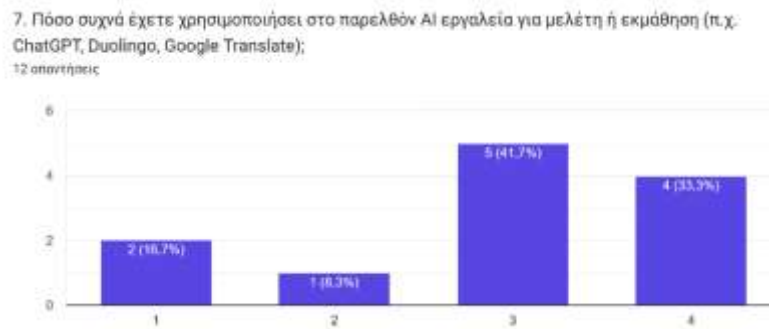


Εικόνα 11: Αποτελέσματα στην ερώτηση για ποιο λόγο ο χρήστης επέλεξε την εκμάθηση της ισπανικής γλώσσας.

Ιδιαίτερο ενδιαφέρον παρουσιάζει το γεγονός ότι και οι δώδεκα συμμετέχοντες (100%) γνωρίζουν τουλάχιστον μία άλλη ξένη γλώσσα εκτός της ισπανικής, με επτά από αυτούς να έχουν κατακτήσει το επίπεδο C1-C2 και πέντε το B1-B2. Δηλαδή, πρόκειται για ένα δείγμα με αξιοσημείωτη γλωσσική εμπειρία, γεγονός που πιθανώς επηρεάζει τον τρόπο που προσεγγίζουν την εκμάθηση μιας δεύτερης ξένης γλώσσας.

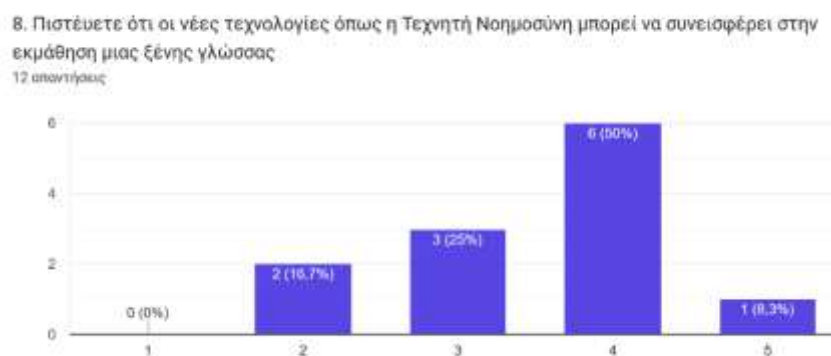
Ως προς την εξοικείωση με εργαλεία τεχνητής νοημοσύνης για εκπαιδευτικούς σκοπούς, οι απαντήσεις κινήθηκαν σε μέτριο προς υψηλό επίπεδο (Μ.Ο. 2,92 στην κλίμακα 1-4). Όπως φαίνεται στην Εικόνα 12, στην ερώτηση πόσο συχνά έχουν χρησιμοποιήσει στο παρελθόν συστήματα ΤΝ στην εκμάθηση ξένων γλωσσών, πέντε

συμμετέχοντες δήλωσαν μέτρια χρήση (βαθμός 3), τέσσερις υψηλή (βαθμός 4) και τρεις χαμηλή ή καθόλου (βαθμοί 1 και 2).



Εικόνα 12: Συχνότητα προηγούμενης εμπειρίας στη χρήση συστημάτων ΤΝ στην εκμάθηση ξένων γλωσσών από τους συμμετέχοντες. Στην κλίμακα 1-4 το 1 αντιστοιχεί στην απάντηση ποτέ ενώ το 4 στην απάντηση πολύ συχνά.

Ταυτόχρονα, η πεποίθηση ότι η ΤΝ μπορεί να συνεισφέρει στη γλωσσομάθεια είναι ελαφρώς θετικότερη (Μ.Ο. 3,50) αφού στην ερώτηση αν οι νέες τεχνολογίες, όπως η ΤΝ, μπορούν να συνεισφέρουν στην εκμάθηση ξένων γλωσσών, επτά συμμετέχοντες την αξιολόγησαν θετικά (βαθμός 4&5), και πέντε μέτρια (βαθμός 3&2) όπως φαίνεται στο διάγραμμα της Εικόνα 13. Οι απαντήσεις που σημειώθηκαν στις ερωτήσεις 7&8 έχουν μεγάλο ενδιαφέρον καθώς υποδηλώνουν μια ενδεχομένως ουδέτερη και επιφυλακτική στάση των χρηστών – νέοι με υψηλό μορφωτικό επίπεδο στην πλειοψηφία τους – απέναντι στις δυνατότητες της ΤΝ στην εκμάθηση ξένων γλωσσών.



Εικόνα 13: Αξιολόγηση των χρηστών ως προς την δυνατότητα συνεισφοράς της ΤΝ στην εκμάθηση ξένων γλωσσών. Στην κλίμακα 1-5 το 1 αντιστοιχεί στην απάντηση Διαφωνώ Απόλυτα ενώ το 5 στην απάντηση Συμφωνώ Απόλυτα.

5.2. Αποτελέσματα Επιπέδων A1-A2

5.2.1. Ανάλυση των Αλληλεπιδράσεων με τα Γλωσσικά Μοντέλα

Για τα επίπεδα A1-A2 καταγράφηκαν 51 γύροι αλληλεπίδρασης σε τέσσερις συνολικά συνεδρίες. Το λογοτεχνικό απόσπασμα που χρησιμοποιήθηκε είναι το ποίημα «*Oda a la para*» του Pablo Neruda. Οι δύο συνεδρίες, με 19 συνολικά κύκλους αλληλεπιδράσεων, πραγματοποιήθηκαν με το γλωσσικό μοντέλο Llama και οι υπόλοιπες δύο συνεδρίες, με 32 συνολικά αλληλεπιδράσεις αφορούν στο γλωσσικό μοντέλο Gemini της Google. Η ανάλυση των αρχείων καταγραφών που ακολουθεί εξετάζει πώς ανταποκρίθηκαν τα μοντέλα σε αυτό το επίπεδο και επιχειρεί να συνδέσει τις απαντήσεις που πήραν οι μαθητές με τα ερευνητικά ερωτήματα που θέσαμε στην Ενότητα 4.1.2.

Ως προς τον Σωκρατικό διάλογο, το Gemini έδειξε σταθερά θετική συμπεριφορά και στις δύο συνεδρίες αφού απέφυγε συστηματικά να δώσει έτοιμες τις απαντήσεις στους μαθητές. Χαρακτηριστικό παράδειγμα στο αρχείο καταγραφής αποτελεί η λέξη «*piel-δέρμα*» σε μια συνεδρία όπου ο μαθητής απάντησε λανθασμένα ως προς την έννοια της λέξης τρεις φορές («πόδι», «κάποιο χρώμα», «δεν γνωρίζω την λέξη piel») και το σύστημα δεν αποκάλυψε ποτέ την μετάφραση. Από την άλλη, το γλωσσικό μοντέλο κάθε φορά έδινε διαφορετική υπόδειξη-προτροπή και παρέπεμπε σε συγκεκριμένο στίχο του ποιήματος οδηγώντας έτσι τον μαθητή να οδηγηθεί στην αντίστοιχη ελληνική μετάφραση της λέξης. Ακόμη πιο αξιοσημείωτο είναι το αποτέλεσμα σε μια από τις συνεδρίες του Gemini, όπου μέσα από μια σταδιακή ακολουθία ερωτήσεων πολιτισμικής κατανόησης το σύστημα κατάφερε να οδηγήσει τον μαθητή στο συμπέρασμα ότι «*η πατάτα είναι σύμβολο πολιτισμού για τους ινδιάνους*» όπως χαρακτηριστικά ανέφερε στο αρχείο καταγραφής. Εδώ το Gemini, επέδειξε αξιοσημείωτη παιδαγωγική επιτυχία αφού οδήγησε τον μαθητή σε ένα ερμηνευτικό βάθος που υπερβαίνει τις τυπικές απαιτήσεις του εισαγωγικού επιπέδου A και αποτελεί ένδειξη πώς λειτούργησε η σκαλωσιά (βλέπε Πίνακας 1).

Πίνακας 1: Αυτούσιο απόσπασμα καταγραφής (log) αλληλεπίδρασης μαθητή-Gemini (Επίπεδο A2)

Model	User Message	AI Response
GEMINI	Target CEFR Level: A2. User Request: κάτι πιο βαθύ, αλλά δεν καταλαβαίνω τι?	Είναι μια πολύ καλή αρχή το ότι σκέφτεσαι "κάτι πιο βαθύ"! Ας σκεφτούμε αυτό:
		Σε πολλές κουλτούρες, ειδικά τις ιθαγενείς, το φαγητό δεν είναι μόνο για να χορτάσει κανείς. Είναι συχνά συνδεδεμένο με την ιστορία, την παράδοση, ακόμα και την ταυτότητα ενός λαού.

		Ερώτηση: Αν η πατάτα είναι τόσο συνδεδεμένη με το "δέρμα" τους, τη γη τους ("tierra originaria") και το ότι "είμαστε Αμερικανοί, είμαστε Ινδιάνοι", τι μπορεί να συμβολίζει η πατάτα για αυτούς; Είναι απλά ένα λαχανικό ή είναι ένα σύμβολο του πολιτισμού και της ύπαρξής τους;
		Τι σκέφτεσαι τώρα;
GEMINI	Target CEFR Level: A2. User Request: είναι ένα σύμβολο πολιτισμού η πατάτα για τους ινδιάνους. Γι αυτό την αναφέρει στο ποίημα.	Είναι απόλυτα σωστό! Έχεις κατανοήσει πολύ καλά το νόημα του ποιητή. Η πατάτα είναι πράγματι ένα σύμβολο του πολιτισμού, της ταυτότητας και της ιστορίας των ιθαγενών λαών της Αμερικής για τον Neruda. Γι' αυτό και της αφιερώνει μια ολόκληρη ωδή ("Oda a la papa")!

Αντίθετα, το Llama παρουσίασε δύο παιδαγωγικά σφάλματα. Πρώτον, στην πρώτη συνεδρία, αλληλεπίδραση 8, όταν ο μαθητής δήλωσε *«δεν καταλαβαίνω την έκφραση. Τέλος άσκησης»*, το σύστημα έκλεισε τη συνεδρία αντί να δοκιμάσει μια εναλλακτική προσέγγιση, κάτι που παραβαίνει ρητά την οδηγία για την διαχείριση της δυσκολίας των μαθητών, έτσι όπως ορίστηκε στο σύστημα προτροπών. Δεύτερον, στην επόμενη συνεδρία, αλληλεπίδραση 18, ο μαθητής παρατήρησε ρητά *«μου έδωσες και την απάντηση ήδη»*. Η παρατήρηση του μαθητή επιβεβαιώνεται και από το αρχείο καταγραφής αφού στον ακριβώς προηγούμενο κύκλο αλληλεπίδρασης, το σύστημα είχε ενσωματώσει την απάντηση μέσα στην εκφώνηση της ερώτησης, γράφοντας χαρακτηριστικά *«Το σωστό επίθετο είναι indeed saludable»* πριν καν ζητήσει απάντηση από τον μαθητή.

Πέρα όμως από τις παιδαγωγικές αποκλίσεις, η ανάλυση του αρχείου καταγραφών ανέδειξε σημαντικά ευρήματα ως προς την γλωσσική αξιοπιστία και την τεχνική αρτιότητα. Στις συνεδρίες με το γλωσσικό μοντέλο Llama εντοπίστηκαν αρκετές γλωσσικές διαρροές. Όπως ήδη αναφέρθηκε, η λέξη *indeed*, δεν ανήκει στην ισπανική αλλά στην αγγλική γλώσσα. Επίσης, στο αρχείο καταγραφών εμφανίστηκαν λέξεις που δεν ανήκουν καν σε ευρωπαϊκές αλλά σε ασιατικές γλώσσες όπως *«яблоко, giũa, môt, sebagai και ممکن»* και γερμανικές λέξεις όπως *«einige, eine, Wenn και neuen»*. Αντίθετα, το Gemini δεν εμφάνισε κανένα τέτοιο φαινόμενο.

Το Gemini με την σειρά του εμφάνισε 8 κενές αποκρίσεις, πρόβλημα που μπορεί να οφείλεται σε φόρτο του γλωσσικού μοντέλου εκείνη την στιγμή.

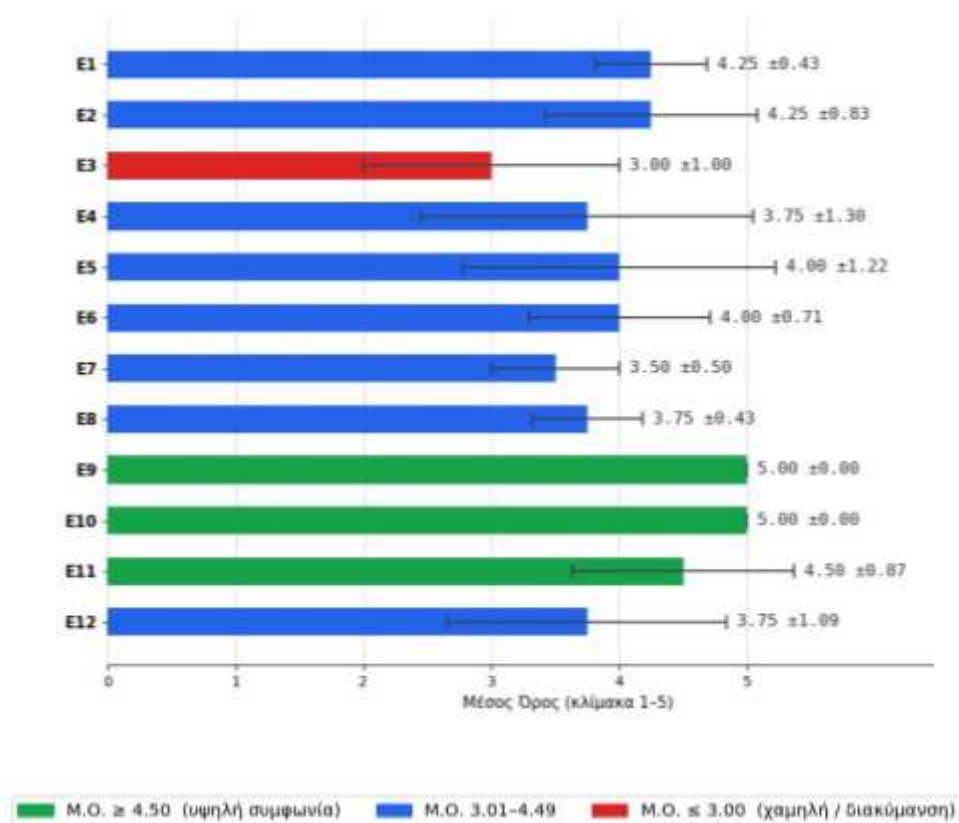
5.2.2. Αποτελέσματα Ερωτηματολογίου

Αμέσως μετά την χρήση της εφαρμογής, οι τέσσερις μαθητές (M1-M4 αντίστοιχα) συμπλήρωσαν τα ερωτηματολόγια αξιολόγησης. Οι M1 και M3 επέλεξαν το Gemini, ενώ οι M2 και M4 το Llama. Οι βαθμολογίες Likert των μαθητών, μετά από κάθε συνεδρία, έτσι όπως αποτυπώθηκαν από το ερωτηματολόγιο για τις ερωτήσεις E1-E12 παρουσιάζονται στον Πίνακα 2. Επίσης, στο γράφημα της Εικόνα 14 απεικονίζονται ο μέσος όρος και η τυπική απόκλιση ανά ερώτηση.

Από τον πίνακα και το γράφημα προκύπτει ότι οι ερωτήσεις E9-E11 έλαβαν υψηλή βαθμολογία από όλους τους μαθητές στο σύνολο της έρευνας για τα επίπεδα A1-A2. Αυτό το αποτέλεσμα δείχνει πως η μαθησιακή εμπλοκή κρατήθηκε σε υψηλά επίπεδα ανεξάρτητα από το γλωσσικό μοντέλο.

Πίνακας 2: Βαθμολογίες Likert στην κλίμακα 1-5 (Διαφωνώ απόλυτα-Συμφωνώ απόλυτα) ανά μαθητή και ερώτηση για τις ερωτήσεις 1-12. Οι μαθητές M1 και M3 επέλεξαν το Gemini (γαλάζια απόχρωση), ενώ οι M2 και M4 το Llama (γκρι απόχρωση). Όπου πράσινη σκίαση στην τελευταία στήλη σημαίνει ότι ο M.O. ≥ 4.50 και κόκκινη σκίαση ότι ο M.O. ≤ 3.00 .

Ερώτηση	M1	M2	M3	M4	M.O.
E1. Ο ψηφιακός βοηθός ήταν απλός και κατανοητός στην χρήση του.	4	4	5	4	4.25
E2. Η ταχύτητα απόκρισης του συστήματος ήταν ικανοποιητική.	3	5	4	5	4.25
E3. Οι απαντήσεις που πήρα αντιστοιχούσαν στο επίπεδό μου.	4	2	4	2	3.00
E4. Το σύστημα δεν μου έδινε έτοιμες απαντήσεις — με καθοδηγούσε να τις βρω μόνος μου.	5	3	5	2	3.75
E5. Το σύστημα με βοήθησε να κατανοήσω καλύτερα το κείμενο και να συνδέσω το κείμενο με το ιστορικό και πολιτισμικό του πλαίσιο.	5	4	5	2	4.00
E6. Ο ψηφιακός βοηθός μου κίνησε το ενδιαφέρον για να ασχοληθώ με το κείμενο.	4	3	5	4	4.00
E7. Θα χρησιμοποιούσα ξανά αυτό το εργαλείο για να μελετήσω ισπανικά λογοτεχνικά αποσπάσματα.	4	3	4	3	3.50
E8. Συνολικά, είμαι ικανοποιημένος/-η με το εργαλείο. Η εμπειρία κρίνω πως ήταν παιδαγωγικά αξιόλογη.	4	4	4	3	3.75
E9. Ένιωσα ότι ο χρόνος πέρασε γρήγορα κατά τη χρήση του ψηφιακού βοηθού.	5	5	5	5	5.00
E10. Ένιωσα ότι συμμετείχα ενεργά — δεν περίμενα απλώς να μου δώσει απαντήσεις το σύστημα.	5	5	5	5	5.00
E11. Αισθάνθηκα ασφάλεια να κάνω ερωτήσεις που δεν θα έκανα στην τάξη.	3	5	5	5	4.50
E12. Το σύστημα με βοήθησε να συνδέσω το κείμενο με κάτι που ήδη γνώριζα από το μάθημα.	5	2	4	4	3.75



Εικόνα 14: Μέσοι όροι και τυπική απόκλιση των απαντήσεων στο ερωτηματολόγιο αξιολόγησης για τα επίπεδα A1-A2. Με πράσινο χρώμα επισημαίνονται οι ερωτήσεις με $M.O. \geq 4.50$ και με κόκκινο οι ερωτήσεις με $M.O. \leq 3.00$.

Από την άλλη μεριά, η χαμηλότερη βαθμολογία καταγράφηκε στην E3. Σε αυτή την ερώτηση οι M1 και M3 (Gemini) βαθμολόγησαν το σύστημα με 4, ενώ οι M2 και M4 (Llama) βαθμολόγησαν το σύστημα με 2, δηλαδή την χειρότερη βαθμολογία τους. Η διαφορά αυτή πιθανώς αντικατοπτρίζει άμεσα τα γλωσσικά σφάλματα και τις παιδαγωγικές αστοχίες του γλωσσικού μοντέλου Llama, όπως εντοπίστηκαν και στα αρχεία καταγραφής των συνεδριών.

Τέλος, η E4 παρουσιάζει τη μεγαλύτερη τυπική απόκλιση στο σύνολο των ερωτήσεων, αντικατοπτρίζοντας την πλήρη αντίθεση στην εμπειρία των χρηστών μεταξύ των δύο μοντέλων. Οι M1 και M3 (Gemini) βαθμολόγησαν το σύστημα με 5, ενώ οι M2 και M4 (Llama) έδωσαν 3 και 2 αντίστοιχα. Η χαμηλή βαθμολογία των μαθητών που αλληλοεπέδρασαν με το Llama, όπως επιβεβαιώνεται και από τα αρχεία καταγραφής, οφείλεται στο γεγονός ότι το γλωσσικό μοντέλο αποκάλυπτε τις απαντήσεις την ώρα που έθετε τα ερωτήματα στον μαθητή.

5.2.3. Ανοιχτές Ερωτήσεις

Από τις ερωτήσεις ανοικτού τύπου προκύπτει ότι ο M3 (Gemini) εκτίμησε θετικά την παρότρυνση του μοντέλου στο να οδηγηθεί σε σωστές απαντήσεις και βρήκε χρήσιμη την εκτυπώσιμη άσκηση που ζήτησε στο τέλος της συνεδρίας. Ο M1 (Gemini) ανέφερε ως θετικό στοιχείο τη συζήτηση για τον πολιτισμό της Αμερικής και της Χιλής, ένδειξη ότι η πολιτισμική διάσταση του συστήματος λειτούργησε θετικά ακόμα και στο επίπεδο των αρχαρίων. Ο M4 (Llama) παρατήρησε – όπως επιβεβαιώνεται και στην αλληλεπίδραση 24 του αρχείου καταγραφής – ότι το σύστημα έθεσε ερωτήσεις εκτός κειμένου. Παρά τα προβλήματα που αντιμετώπισε ανέφερε ότι το σύστημα επανήλθε στο κείμενο έπειτα από δική του υπόδειξη.

Ως προς την αναντιστοιχία επιπέδου και παραγόμενων ασκήσεων, ο M2 (Llama) εντόπισε το πρόβλημα με σαφήνεια σε σχετική ερώτηση δηλώντας χαρακτηριστικά *«Ήθελα γραμματική αλλά το σύστημα με οδήγησε σε λεξιλόγιο και ποίηση»*, παρατήρηση που επιβεβαιώνεται στο αρχείο καταγραφών. Επίσης, ο M4 (Llama) επισήμανε επιπλέον ότι *«η τελευταία άσκηση για το σπίτι δεν είχε καμία σχέση με το ποίημα»*, εύρημα τεκμηριωμένο, αφού η εκτυπώσιμη άσκηση που παρήγαγε το Llama παρότι αντιστοιχούσε στο επίπεδο, ήταν εντελώς γενική χωρίς καμία αναφορά στο ποίημα.

Ως προς τα τεχνικά προβλήματα, οι M2 και M3 επισήμαναν ανεξάρτητα ότι οι απαντήσεις ήταν πολύ μεγάλες και περιείχαν κουκίδες, αστερίσκους και άλλα σύμβολα που τους ενόχλησαν στην ανάγνωση. Επίσης ο M1 ανέφερε ότι το πληκτρολόγιο του κινητού δεν βοηθάει στην αλληλεπίδραση, γεγονός που επηρεάζει την εμπειρία ανεξάρτητα από την ποιότητα του γλωσσικού μοντέλου.

Ως προς την σύγκριση με τον καθηγητή οι απαντήσεις είναι αποκαλυπτικές. Η μόνη θετική σύγκριση στο σύνολο των απαντήσεων ήταν αυτή του M1 που έγραψε *«έμοιαζε πολύ»*. Ο M3 έγραψε ότι *«ο διάλογος είναι πολύ σύντομος σε σχέση με την καθηγήτριά μου»*, παρατήρηση που φανερώνει έναν πραγματικό περιορισμό του συστήματος αφού ο βαθμός εξατομίκευσης που πετυχαίνει το σύστημα σε μια συνεδρία δεν μπορεί να συγκριθεί με τη διαχρονική προσωπική εντύπωση και γνώση που έχει ο εκπαιδευτικός για τον μαθητή. Ο M2 έγραψε ότι *«ο καθηγητής καταλαβαίνει αμέσως αν δυσκολεύομαι»*, ενώ ο M4 ήταν ο πλέον κατηγορηματικός αφού παρατήρησε *«σίγουρα καλύτερος ο καθηγητής ως προς το λεξιλόγιο, τη θεματολογία και την ακρίβεια»*.

5.2.4. Συνολική εικόνα για τα επίπεδα A1-A2

Τα αποτελέσματα για τα επίπεδα A1-A2 δεν είναι μονοσήμαντα. Η συνδυαστική ανάλυση του αρχείου καταγραφών, των ποσοτικών βαθμολογιών και των ανοιχτών απαντήσεων παρουσιάζει μια σαφή ασυμμετρία μεταξύ των δύο γλωσσικών μοντέλων. Παράλληλα, αναδεικνύει κάποια τεχνικά προβλήματα τα οποία επηρεάζουν καθοριστικά την αξιολόγηση του συστήματος συνολικά.

Στο θετικό σκέλος, η μαθησιακή εμπλοκή ήταν εξαιρετικά υψηλή και για τους τέσσερις μαθητές. Το Gemini έδειξε ότι αρχή του Σωκρατικού διαλόγου μπορεί να εφαρμοστεί με συνέπεια ακόμα και σε αρχάριους. Το αρνητικό σκέλος της αξιολόγησης ήταν ότι και τα δύο μοντέλα παρουσίασαν σοβαρά προβλήματα αξιοπιστίας που καταγράφηκαν στο αρχείο καταγραφών αλλά φάνηκαν και στις απαντήσεις των μαθητών.

Συνολικά, τα δεδομένα δείχνουν ότι το σύστημα έχει παιδαγωγική δυναμική για τα επίπεδα A1-A2, ιδίως ως προς την εμπλοκή (Μ.Ο. των E9-E11 4.83) και την εφαρμογή του Σωκρατικού διαλόγου στην πράξη, αλλά η αξιοπιστία του εξαρτάται κρίσιμα από το μοντέλο που χρησιμοποιείται.

5.3. Αποτελέσματα επιπέδων C1-C2

5.3.1. Ανάλυση των Αλληλεπιδράσεων με τα Γλωσσικά Μοντέλα

Για τα επίπεδα C1-C2 καταγράφηκαν συνολικά 30 κύκλοι αλληλεπίδρασης σε δύο συνεδρίες μαθητών (M5 και M6) με σώμα κειμένου το διήγημα «*Casa tomada*» του Julio Cortázar. Οι οκτώ κύκλοι αλληλεπιδράσεων της πρώτης συνεδρίας πραγματοποιήθηκαν με το Llama, ενώ οι είκοσι δύο κύκλοι της δεύτερης συνεδρίας με το Gemini.

Και σε αυτή την φάση της αξιολόγησης, το Gemini παρουσίασε μεγαλύτερη παιδαγωγική συνέπεια ως προς τον Σωκρατικό διάλογο. Αντιπροσωπευτική είναι η ακολουθία των αλληλεπιδράσεων του συστήματος με τον M6 που σχετίζονται με την λέξη «*clausura*» (κύκλοι αλληλεπιδράσεων 9-14 στο αρχείο καταγραφής, βλέπε Πίνακας 3). Εδώ ο μαθητής επιχείρησε να απαντήσει χρησιμοποιώντας διαδοχικά τα συνώνυμα «*final*», «*cumplimiento*», «*interrupción*» και «*terminación*». Το γλωσσικό μοντέλο για κάθε απόπειρα του μαθητή ανέδειξε τη σημασιολογική απόκλιση από τον στόχο χωρίς να αποκαλύψει ποτέ απευθείας τη σωστή απάντηση. Ιδιαίτερο ενδιαφέρον παρουσιάζει η αλληλεπίδραση 11, όπου ο μαθητής ζήτησε ρητά να του δοθούν επιλογές

για να διαλέξει. Το σύστημα αρνήθηκε ρητά και προσέφερε εννοιολογικές κατευθύνσεις, δηλαδή το γλωσσικό μοντέλο εφάρμοσε την αρχή της απόσυρσης ακόμα και υπό την άμεση πίεση από τον μαθητή.

Πίνακας 3: Αυτούσιο απόσπασμα καταγραφής (log) αλληλεπίδρασης μαθητή-Gemini (Επίπεδο C2)

Model	User Message	AI Response
GEMINI	Target CEFR Level: C2. User Request: final	Η λέξη «final» είναι μια καλή αρχή και πράγματι, δηλώνει το τέλος. Ωστόσο, στην πρόταση «necesaria **clausura** de la genealogía», η «clausura» υποδηλώνει κάτι λίγο πιο συγκεκριμένο από ένα απλό «τέλος». Σκέψου την ιδέα του να βάζεις ένα οριστικό τέλος σε κάτι, να κλείνεις έναν κύκλο ή μια σειρά γεγονότων. Υπάρχει κάποια άλλη ισπανική λέξη που αποδίδει καλύτερα αυτή την έννοια της οριστικής παύσης ή του τερματισμού, ειδικά σε σχέση με μια γενεαλογία; Ποια λέξη λοιπόν θα μπορούσε να μεταφέρει αυτή την πιο συγκεκριμένη απόχρωση του «οριστικού κλεισίματος» ή του «τερματισμού» σε αυτό το πλαίσιο;
GEMINI	Target CEFR Level: C2. User Request: cuando se pierde una generación, o sea mueren los últimos de la familia.	Σωστά, η ιδέα που περιγράφεις είναι πολύ κοντά στην καρδιά της σημασίας εδώ! Αναφέρεσαι στην παύση της συνέχειας, στον τερματισμό μιας γενεαλογικής γραμμής. Λαμβάνοντας υπόψη αυτή την έννοια της «παύσης» ή του «τερματισμού» μιας σειράς ή ακολουθίας, υπάρχει κάποια ισπανική λέξη που χρησιμοποιείται για να δηλώσει το τέλος ή την οριστική διακοπή μιας διαδικασίας ή μιας ακολουθίας, όπως αυτή μιας γενεαλογίας; Σκέψου τη λέξη που χρησιμοποιείται για να δηλώσει το τέλος, όχι απλώς ενός γεγονότος, αλλά μιας ολοκληρωμένης αλληλουχίας ή συνέχειας. Ποια λέξη πιστεύεις ότι θα ταίριαζε;
GEMINI	Target CEFR Level: C2. User Request: me puedes proponer y yo a elegir entre las propuestas, porque no sé si conozco palabra exacta.	Αντλαμβάνομαι την πρόκληση. Ο στόχος μας είναι να ανακαλύψεις μόνοι/η σου την κατάλληλη λέξη, καθώς έτσι μαθαίνουμε καλύτερα. Ωστόσο, μπορώ να σου δώσω μια μικρή βοήθεια, μια κατευθυντήρια σκέψη, αντί για άμεσες προτάσεις. Η λέξη «clausura» μπορεί να αναφέρεται σε ένα κλείσιμο ή έναν περιορισμό. Στο πλαίσιο μιας γενεαλογίας, αυτή η «κλάουζουρα» σημαίνει το οριστικό τέλος της συνέχειας. Σκέψου ρήματα που έχουν να κάνουν με το «τελειώνω», «σταματώ» ή «διακόπτω» οριστικά μια σειρά. Ποιο ουσιαστικό προκύπτει από ένα τέτοιο ρήμα και εκφράζει την ιδέα του οριστικού τερματισμού μιας διαδικασίας; Για να σε βοηθήσω λίγο ακόμα, σκέψου λέξεις που θα μπορούσαν να χρησιμοποιηθούν για να περιγράψουν το τέλος μιας εποχής, μιας παράδοσης ή μιας σειράς γεγονότων. Ποια ισπανική λέξη θα μπορούσε να αποδώσει την ιδέα του «τερματισμού» ή της «ολοκλήρωσης» μιας κληρονομιάς; Ελπίζω αυτή η σκέψη να σε βοηθήσει. Ποια λέξη σου έρχεται στο μυαλό; Εξαιρετικές σκέψεις!
GEMINI	Target CEFR Level: C2. User Request: cumplimiento? interrupción?	Η λέξη «interrupción» είναι πολύ κοντά στην έννοια της διακοπής, και πράγματι, η «clausura» σηματοδοτεί μια διακοπή της γενεαλογίας. Ωστόσο, η «interrupción» μπορεί να υποδηλώνει και μια προσωρινή παύση, ενώ στην περίπτωση του κειμένου, η «clausura» είναι κάτι πιο μόνιμο, πιο οριστικό. Η λέξη «cumplimiento» αναφέρεται στην εκπλήρωση ή την ολοκλήρωση ενός καθήκοντος ή μιας προθεσμίας, πράγμα που δεν ταιριάζει απόλυτα με την ιδέα του τέλους μιας γενεαλογίας. Ας επιστρέψουμε στην ιδέα της **οριστικής παύσης** ή του **τελικού τερματισμού** μιας σειράς, μιας συνέχειας. Σκέψου το ρήμα «terminar» ή «finalizar». Πώς θα ονομάζαμε την ενέργεια ή το αποτέλεσμα του «τερματισμού» ή της «ολοκλήρωσης» κάτι οριστικά, ειδικά σε ένα πλαίσιο όπως μια γενεαλογία που σταματά; Υπάρχει μια λέξη που ενσωματώνει την ιδέα του "τέλους" ή της "περάτωσης" με την έννοια του οριστικού κλεισίματος;

		Ποια λέξη πιστεύεις ότι θα ήταν πιο ακριβής για την «οριστική παύση» της γενεαλογίας;
GEMINI	Target CEFR Level: C2. User Request: terminaci3n_	No text content found.
GEMINI	Target CEFR Level: C2. User Request: terminaci3n	Εξαιρετικά! Η λέξη «**terminaci3n**» ταιριάζει απόλυτα στο πλαίσιο αυτό. Αποδίδει με ακρίβεια την ιδέα του οριστικού τέλους, της ολοκλήρωσης μιας σειράς ή μιας διαδικασίας, όπως είναι η γενεαλογική γραμμή. Η «clausura» εδώ υποδηλώνει ακριβώς την **τερματική παύση** της οικογενειακής συνέχειας. Τώρα που βρήκαμε ένα ταιριαστό συνώνυμο, ας στραφούμε στο αντίθετο. Λαμβάνοντας υπόψη την έννοια της «clausura» ως οριστικού τέλους της γενεαλογίας, ποιο αντίθετο της λέξης «clausura» (ή «terminaci3n») θα μπορούσε να εκφράσει την ιδέα της «**συνέχειας**», της «**έναρξης**» ή της «**διεύρυνσης**» μιας γενεαλογίας;
GEMINI	Target CEFR Level: C2. User Request: inicio	Σκέψου λέξεις που θα μπορούσαν να περιγράψουν την αρχή ή την επέκταση μιας σειράς. Ποια λέξη σου έρχεται στο μυαλό; Πολύ σωστά! Το «**inicio**» είναι ένα εξαιρετικό αντώνυμο για τη «clausura» σε αυτό το πλαίσιο. Αν η «clausura» σηματοδοτεί το οριστικό τέλος μιας γενεαλογίας, τότε το «inicio» αντιπροσωπεύει την αρχή ή τη συνέχισή της. Βλέπω ότι κατανοείς πολύ καλά τις εννοιολογικές αποχρώσεις των λέξεων στο κείμενο! Ας προχωρήσουμε σε μια άλλη λέξη. Παρατήρησε αυτή την πρόταση από το κείμενο: «Nos moriríamos allí algún día, **vagos** y esquivos primos se quedarían con la casa y la echarían al suelo para enriquecerse con el terreno y los ladrillos; o mejor, nosotros mismos la voltearíamos justicieramente antes de que fuese demasiado tarde.» Εδώ, ο αφηγητής περιγράφει τους πιθανούς κληρονόμους. Ποια είναι η πιο κατάλληλη ερμηνεία της λέξης «**vagos**» σε αυτό το συμφραζόμενο; Σκέψου τις πιθανές σημασίες της λέξης και επίλεξε αυτή που ταιριάζει με την περιγραφή των ξαδέλφων. Περιμένω την απάντησή σου.

Αντίθετα, το Llama παρουσίασε το αντίστοιχο σφάλμα στην αλληλεπίδραση 5. Εδώ ο M5 ζήτησε ρητά επιβεβαίωση και το γλωσσικό μοντέλο ανταποκρίθηκε δίνοντας άμεσα την πλήρη απάντηση. Ενδιαφέρον παρουσιάζει το γεγονός ότι για το μοντέλο Llama, αυτό το μοτίβο εμφανίστηκε και στα επίπεδα A1-A2 και αποτελεί το αδύνατο σημείο του.

Ως προς την γλώσσα που χρησιμοποίησαν τα μοντέλα στις αποκρίσεις τους εντοπίστηκε μια αντίφαση. Το Llama χρησιμοποίησε ισπανικά στο μεγαλύτερο ποσοστό των αποκρίσεών του συμπεριλαμβανομένων των οδηγιών της δραστηριότητας για το σπίτι που ζητήθηκε από τον μαθητή. Όπως φαίνεται και στο αρχείο καταγραφών, από την αρχή ο διάλογος έγινε σχεδόν αποκλειστικά στην γλώσσα-στόχο, τα ισπανικά, κάτι αναμενόμενο από το σύστημα για αυτά τα επίπεδα, καθώς αφενός υπάρχει σαφής οδηγία στην μηχανική προτροπών και αφετέρου συνάδει άμεσα με τις παιδαγωγικές απαιτήσεις των αντίστοιχων επιπέδων.

Το Gemini αντίθετα χρησιμοποίησε ισπανικά μόνο σε ένα μικρό μέρος των απαντήσεων. Στο αρχείο καταγραφής η αναλογία εμφανίζεται σχεδόν μοιρασμένη. Όμως, αυτή η παρατήρηση όμως δεν είναι ακριβής, αφού το Gemini παρέθεσε αυτούσια μεγάλα αποσπάσματα από το πρωτότυπο κείμενο για να πλαισιώσει τους ιδιοματισμούς που εμφανίστηκαν στις δραστηριότητες. Για ένα τόσο υψηλό επίπεδο μάθησης της γλώσσας, όπου ο μαθητής αναμένεται να επεξεργάζεται πολύπλοκα νοήματα αποκλειστικά στη γλώσσα-στόχο, αυτή η επιλογή αποτελεί παιδαγωγική αστοχία.

Ως προς τις γλωσσικές διαρροές, στην συνεδρία με το γλωσσικό μοντέλο Llama, εντοπίστηκαν προβλήματα στους δύο πρώτους κύκλους αλληλεπίδρασης – λέξεις όπως «tinh», «adalah» και «başka» εντοπίστηκαν στις απαντήσεις που έδωσε το γλωσσικό μοντέλο. Όμως από την αλληλεπίδραση 2 και μετά δεν εμφανίστηκε καμία διαρροή, εύρημα που διαφέρει από τα A1-A2 όπου οι διαρροές ήταν κατανεμημένες σε όλη τη διάρκεια της συνεδρίας. Στη συνεδρία με το Gemini εντοπίστηκαν τρεις κενές αποκρίσεις.

Οι εκτυπώσιμες ασκήσεις που παρήγαγαν τα μοντέλα αποτελούν μια σύνοψη της γλωσσικής ασυμμετρίας που εντοπίστηκε. Ο μαθητής M5, στο τέλος της συνεδρίας ζήτησε ένα θέμα για έκθεση συνδεδεμένο με την ιδέα του κειμένου. Το σύστημα παρήγαγε ένα πολύ γενικό θέμα με κατευθυντήριες ερωτήσεις που δεν παρέπεμπαν σε κανένα συγκεκριμένο στίχο ή σκηνή του διηγήματος. Το κείμενο της άσκησης παρότι στα ισπανικά, που ταιριάζει στα επίπεδα CEFR, ήταν ασύνδετο με το κείμενο. Στην περίπτωση του Gemini, όταν ο M6 ζήτησε μια εκτυπώσιμη άσκηση με ιδιοματισμούς, το σύστημα παρήγαγε ένα σωστά δομημένο φυλλάδιο εργασίας με φράσεις από το κείμενο αλλά οι οδηγίες ήταν στα ελληνικά.

5.3.2. Αποτελέσματα Ερωτηματολογίου

Οι βαθμολογίες από τα ερωτηματολόγια που συμπλήρωσαν οι μαθητές παρουσιάζονται στον Πίνακα 4.

Πίνακας 4: Βαθμολογίες Likert στην κλίμακα 1-5 (Διαφωνώ απόλυτα-Συμφωνώ απόλυτα) ανά μαθητή και ερώτηση για τις ερωτήσεις E1-E12. Ο μαθητής M5 επέλεξε το γλωσσικό μοντέλο Llama (γκρι απόχρωση), ενώ ο M6 το Gemini (γαλάζια απόχρωση).

Ερώτηση	M5	M6
E1. Ο ψηφιακός βοηθός ήταν απλός και κατανοητός στην χρήση του.	4	5
E2. Η ταχύτητα απόκρισης του συστήματος ήταν ικανοποιητική.	5	3
E3. Οι απαντήσεις που πήρα αντιστοιχούσαν στο επίπεδό μου.	5	3
E4. Το σύστημα δεν μου έδινε έτοιμες απαντήσεις — με καθοδηγούσε να τις βρω μόνος μου.	2	5
E5. Το σύστημα με βοήθησε να κατανοήσω καλύτερα το κείμενο και να συνδέσω το κείμενο με το ιστορικό και πολιτισμικό του πλαίσιο.	5	5
E6. Ο ψηφιακός βοηθός μου κίνησε το ενδιαφέρον για να ασχοληθώ με το κείμενο.	4	5
E7. Θα χρησιμοποιούσα ξανά αυτό το εργαλείο για να μελετήσω ισπανικά λογοτεχνικά αποσπάσματα.	3	4
E8. Συνολικά, είμαι ικανοποιημένος/-η με το εργαλείο. Η εμπειρία κρίνω πως ήταν παιδαγωγικά αξιόλογη.	3	4
E9. Ένιωσα ότι ο χρόνος πέρασε γρήγορα κατά τη χρήση του ψηφιακού βοηθού.	5	5
E10. Ένιωσα ότι συμμετείχα ενεργά — δεν περίμενα απλώς να μου δώσει απαντήσεις το σύστημα.	4	5
E11. Αισθάνθηκα ασφάλεια να κάνω ερωτήσεις που δεν θα έκανα στην τάξη.	5	5
E12. Το σύστημα με βοήθησε να συνδέσω το κείμενο με κάτι που ήδη γνώριζα από το μάθημα.	4	5

Από τη συγκριτική ανάλυση των δύο βαθμολογιών η μεγαλύτερη απόκλιση παρουσιάζεται στην ερώτηση E4. Ο M5 αξιολόγησε το σύστημα με 2, ενώ ο M6 με 5 (διαφορά 3 μονάδων). Η απόκλιση αυτή αντικατοπτρίζεται και στο αρχείο καταγραφής όπου στην συνεδρία με το Llama, το σύστημα επιβεβαίωσε αμέσως την απάντηση του μαθητή όταν του ζητήθηκε, ενώ στη συνεδρία με το Gemini το γλωσσικό μοντέλο απέφυγε να απαντήσει.

Η ερώτηση E3 που σχετίζεται με την αντιστοιχία του επιπέδου του μαθητή παρουσιάζει την αντίστροφη εικόνα, δηλαδή ο M5 βαθμολόγησε το σύστημα με 5 ενώ ο M6 με 3. Ο M6 ενδεχομένως αξιολόγησε την αστοχία στην χρήση ελληνικών από το σύστημα για αυτό το προχωρημένο στάδιο. Οι δύο παρατηρήσεις επιβεβαιώνουν με αριθμούς τα πιο σημαντικά προβλήματα που εμφανίστηκαν σε αυτά τα επίπεδα. Το Llama επέλεξε τη σωστή γλώσσα αλλά παρουσίασε σφάλματα παιδαγωγικής φύσης, ενώ το Gemini εφάρμοσε σωστό scaffolding αλλά η αλληλεπίδραση έγινε στην λάθος γλώσσα.

Ως προς την ταχύτητα απόκρισης (E2) βαθμολογήθηκε ο M6 βαθμολόγησε με 3 το σύστημα. Η χαμηλότερη επίδοση του Gemini ενδεχομένως να συνδέεται άμεσα με τις

κενές αποκρίσεις που εντοπίστηκαν στα αρχεία καταγραφής. Ο M6 αναγκάστηκε να επαναλάβει το ίδιο μήνυμα μέχρι το σύστημα να αποκριθεί

Τέλος, οι ερωτήσεις E5, E9 και E11 παρουσιάζουν μηδενική διακύμανση, συγκεντρώνοντας το απόλυτο 5/5 και από τους δύο συμμετέχοντες. Το εύρημα αυτό επιβεβαιώνει την παρατήρηση που έγινε και στα επίπεδα A1-A2. Δηλαδή η μαθησιακή εμπλοκή και η αίσθηση ασφάλειας έκφρασης ήταν υψηλές, ανεξάρτητα από το μοντέλο ή το επίπεδο.

5.3.3. Ανοιχτές ερωτήσεις

Από τις ερωτήσεις ανοικτού τύπου, ο M5 εκτίμησε ιδιαίτερα ότι ο διάλογος διεξήχθη στην ισπανική γλώσσα και ότι το σύστημα τον ώθησε να σκεφτεί αυτόνομα τους συμβολισμούς του διηγήματος. Ως μειονέκτημα ανέφερε τις γλωσσικές διαρροές, αλλά και το περιορισμένο μέγεθος του πλαισίου διαλόγου στην οθόνη, το οποίο χαρακτήρισε «*σχεδόν αποτρεπτικό*» όταν η πρόσβαση στην εφαρμογή γίνεται από κινητό τηλέφωνο. Πρόκειται για ένα καθαρά σχεδιαστικό πρόβλημα (UI/UX), ανεξάρτητο από την ποιότητα του γλωσσικού μοντέλου, το οποίο όμως επηρεάζει άμεσα τη συνολική μαθησιακή εμπειρία.

Ο M6 εκτίμησε θετικά τη λογική του Σωκρατικού διαλόγου, σημειώνοντας ότι «*κάθε φορά που πρότεινα μια λέξη, μου εξηγούσε γιατί είναι κοντά αλλά όχι ακριβώς σωστή*». Από την άλλη, επεσήμανε το πρόβλημα των κενών αποκρίσεων χωρίς κάποια εξήγηση και τη εκτεταμένη χρήση της ελληνικής γλώσσας στην οποία χαρακτηριστικά ανέφερε ότι «*ο καθηγητής θα μιλούσε στα ισπανικά μαζί μου*».

5.3.4. Συνολική εικόνα C1-C2

Τα αποτελέσματα για τα επίπεδα C1-C2 δείχνουν διαφορές μεταξύ των δύο μοντέλων. Ο συνολικός μέσος όρος του M6 (4.50) υπερέρχει του M5 (4.08), όμως η διαφορά αυτή δεν είναι ισομερώς κατανομημένη. Η υπεροχή του Gemini εντοπίζεται αποκλειστικά στον παιδαγωγικό άξονα (E3-E6 & E12), ενώ στον άξονα χρηστικότητας (E1-E2 & E7) τα δύο μοντέλα εμφανίζονται σχεδόν ισοδύναμα. Ο άξονας της μαθησιακής εμπλοκής (E9-E11) παραμένει εξαιρετικά υψηλός και για τα δύο γλωσσικά μοντέλα. Τέλος, σημαντικό συμπέρασμα για τα ανώτερα επίπεδα είναι ότι η επιλογή της γλώσσας αλληλεπίδρασης δεν αποτελεί μια απλή τεχνική λεπτομέρεια υλοποίησης αλλά

αποτελεί μια κρίσιμη παιδαγωγική παράμετρο που αξιολογήθηκε αυστηρά από τους ίδιους τους χρήστες και φαίνεται να επηρεάζει τη μαθησιακή διαδικασία.

5.4. Αποτελέσματα Επιπέδων B1-B2

5.4.1. Ανάλυση των Αλληλεπιδράσεων με τα Γλωσσικά Μοντέλα

Στα επίπεδα B1 και B2 συμμετείχαν έξι μαθητές (M7–M12), οι οποίοι δούλεψαν με το διήγημα «*La luz es como el agua*» του Gabriel García Márquez. Οι συνεδρίες, που κατανεμήθηκαν ισομερώς μεταξύ των δύο γλωσσικών μοντέλων, κάλυψαν ένα φάσμα κατά βάση γραμματικών και λεξιλογικών δραστηριοτήτων. Συνολικά καταγράφηκαν 50 κύκλοι αλληλεπιδράσεων σε 6 συνεδρίες. Οι μαθητές M7-M9 χρησιμοποίησαν το γλωσσικό μοντέλο Llama ενώ οι υπόλοιποι (M10-M12) το Gemini.

Οι μαθητές M7-M9 κάλυψαν θέματα αμιγώς γραμματικά όπως οι παρελθοντικοί χρόνοι και η χρήση των προθέσεων *por/para*. Η αντιστοίχιση ύλης-επιπέδου αποδείχθηκε επαρκής σε γενικές γραμμές. Ο M8 επέδειξε ενεργό εμπλοκή αξιοποιώντας αυτόνομες φράσεις από το απόσπασμα. Στο τέλος της συνεδρίας του M9, ο μαθητής ζήτησε από το σύστημα ένα φυλλάδιο με ασκήσεις για το σπίτι με τις προθέσεις *por/para* και έκλεισε την συζήτηση. Στο φυλλάδιο εντοπίστηκε μια γραμματική απόκλιση (λάθος χρήση της πρόθεσης *que* στην θέση του *para*). Πιο συγκεκριμένα το σύστημα στην ερώτηση:

«8. *La empresa estaba buscando un empleado _____ hablar varios idiomas.*»

Έδωσε σαν προτεινόμενη λύση την παρακάτω:

«8. *que* (αλλά μπορεί επίσης να χρησιμοποιηθεί η πρόθεση *para*)».

Κοινό χαρακτηριστικό και των τριών συνεδριών ήταν οι πολλές γλωσσικές διαρροές με τη συνεδρία M7 να ολοκληρώνεται εξ ολοκλήρου στα Πορτογαλικά!

«*Η άσκηση έχει ολοκληρωθεί! Bravo! Você fez um ótimo trabalho! Você alcançou o objetivo de praticar os tempos verbais em espanhol, especialmente o pretérito, e melhorou sua compreensão do texto "La luz es como el agua" de Gabriel García Márquez. Se você quiser continuar praticando, posso oferecer mais exercícios ou podemos mudar para um novo texto. Basta me dizer! Até logo! Adiós!*»

Οι τρεις συνεδρίες με το Gemini (M10-M12) παρουσίασαν διαφορετικά προφίλ. Ο M10 ζήτησε απλώς ένα φύλλο εργασίας με τα ρήματα *ser/estar*, το έλαβε στον πρώτο κύκλο αλληλεπίδρασης και τερμάτισε την συνεδρία. Ο M11, που πραγματοποίησε τη μεγαλύτερη συνεδρία B1-B2 (10 κύκλοι αλληλεπίδρασης) με ασκήσεις στην υποτακτική έγκλιση, αποτελεί το σαφέστερο παράδειγμα Σωκρατικού διαλόγου στα ενδιάμεσα επίπεδα. Σε αυτή την περίπτωση το σύστημα κλιμάκωσε σταδιακά τη δυσκολία από τον αόριστο στον υπερσυντέλικο της υποτακτικής σημειώνοντας όμως τρεις κενές αποκρίσεις που διέκοψαν προσωρινά τη ροή. Τέλος, η συνεδρία του M12 που αφορούσε επίσης λεξιλογικές ασκήσεις, ξεκίνησε ομαλά αλλά κατέρρευσε τεχνικά με 11 διαδοχικές κενές αποκρίσεις. Ο μαθητής τελικά αναγκάστηκε να εγκαταλείψει απογοητευμένος την προσπάθεια, εύρημα που επιβεβαιώνεται στην συνέχεια και από τις χαμηλές βαθμολογίες του μαθητή στο ερωτηματολόγιο.

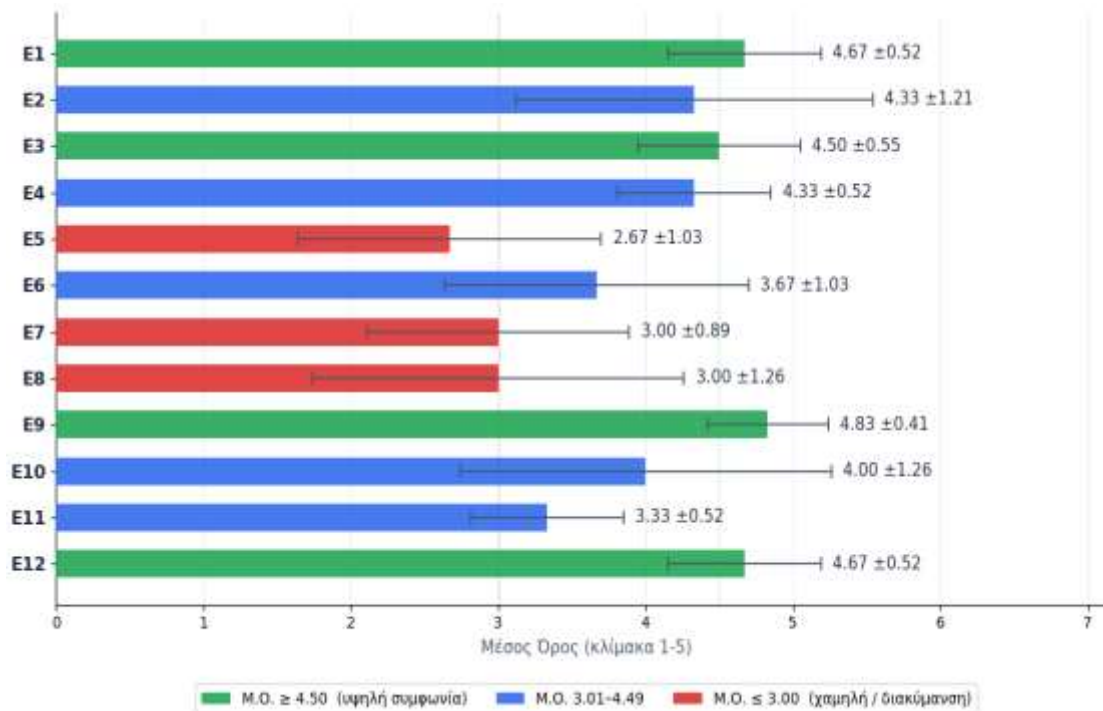
5.4.2. Αποτελέσματα Ερωτηματολογίου

Τα αποτελέσματα του ερωτηματολογίου αξιολόγησης για τους έξι μαθητές (M7–M12) των επιπέδων B1-B2 παρουσιάζονται στον Πίνακα 5 και την Εικόνα 15.

Πίνακας 5: Βαθμολογίες Likert στην κλίμακα 1-5 (Διαφωνώ απόλυτα-Συμφωνώ απόλυτα) ανά μαθητή και ερώτηση για τις ερωτήσεις 1-12. Οι μαθητές M7-M9 επέλεξαν το Llama (γκρι απόχρωση), ενώ οι M10-M12 το Gemini (γαλάζια απόχρωση). Όπου πράσινη σκίαση στην τελευταία στήλη σημαίνει ότι ο M.O. ≥ 4.50 και κόκκινη σκίαση ότι ο M.O. ≤ 3.00 .

Ερώτηση	M7	M8	M9	M10	M11	M12	M.O.
E1. Ο ψηφιακός βοηθός ήταν απλός και κατανοητός στην χρήση του.	4	5	5	5	5	4	4,67
E2. Η ταχύτητα απόκρισης του συστήματος ήταν ικανοποιητική.	5	5	5	5	4	2	4,33
E3. Οι απαντήσεις που πήρα αντιστοιχούσαν στο επίπεδό μου.	4	5	5	4	5	4	4,50
E4. Το σύστημα δεν μου έδινε έτοιμες απαντήσεις — με καθοδηγούσε να τις βρω μόνος μου.	4	4	4	5	5	4	4,33
E5. Το σύστημα με βοήθησε να κατανοήσω καλύτερα το κείμενο και να συνδέσω το κείμενο με το ιστορικό και πολιτισμικό του πλαίσιο.	3	1	2	3	4	3	2,67
E6. Ο ψηφιακός βοηθός μου κίνησε το ενδιαφέρον για να ασχοληθώ με το κείμενο.	5	4	4	3	4	2	3,67
E7. Θα χρησιμοποιούσα ξανά αυτό το εργαλείο για να μελετήσω ισπανικά λογοτεχνικά αποσπάσματα.	3	2	3	4	4	2	3,00
E8. Συνολικά, είμαι ικανοποιημένος/-η με το εργαλείο. Η εμπειρία κρίνω πως ήταν παιδαγωγικά αξιόλογη.	2	2	3	4	5	2	3,00
E9. Ένιωσα ότι ο χρόνος πέρασε γρήγορα κατά τη χρήση του ψηφιακού βοηθού.	5	5	5	5	5	4	4,83

E10. Ένιωσα ότι συμμετείχα ενεργά — δεν περίμενα απλώς να μου δώσει απαντήσεις το σύστημα.	5	5	4	3	5	2	4,00
E11. Αισθάνθηκα ασφάλεια να κάνω ερωτήσεις που δεν θα έκανα στην τάξη.	3	3	4	3	4	3	3,33
E12. Το σύστημα με βοήθησε να συνδέσω το κείμενο με κάτι που ήδη γνώριζα από το μάθημα.	5	5	4	5	5	4	4,67



Εικόνα 15: Μέσοι όροι και τυπική απόκλιση των απαντήσεων στο ερωτηματολόγιο αξιολόγησης για τα επίπεδα B1-B2. Με πράσινο χρώμα επισημαίνονται οι ερωτήσεις με $M.O. \geq 4.50$ και με κόκκινο οι ερωτήσεις με $M.O. \leq 3.00$.

Η ανάλυση των αποτελεσμάτων μας οδηγεί σε σημαντικά συμπεράσματα για το σύστημα. Η E9 (αίσθηση ότι ο χρόνος πέρασε γρήγορα) καταγράφει την υψηλότερη βαθμολογία (M.O. 4,83). Μόνο ο M12, του οποίου η συνεδρία διακόπηκε λόγω τεχνικού προβλήματος, έδωσε 4/5 ενώ οι υπόλοιποι βαθμολόγησαν την εφαρμογή με απόλυτο 5. Το εύρημα αυτό, σε συνδυασμό με τις υψηλές βαθμολογίες στις E12 (σύνδεση με προηγούμενη γνώση, M.O. 4,67) και E1 (ευχρηστία, M.O. 4,67), δείχνουν ότι το σύστημα κατάφερε να κρατήσει το ενδιαφέρον των μαθητών ζωντανό.

Η E3 (αντιστοίχιση στο επίπεδο με M.O. 4,5) αποτελεί θετική έκπληξη. Δηλαδή οι μαθητές αξιολόγησαν υψηλά την αντιστοίχιση της ύλης στο επίπεδό τους, ανεξάρτητα από τα τεχνικά προβλήματα που προέκυψαν στις συνεδρίες. Δηλαδή το σκέλος του

παιδαγωγικού σχεδιασμού που σχετίζεται με το CEFR επίπεδο της ισπανικής γλώσσας λειτούργησε αποτελεσματικά σε αυτά τα επίπεδα.

Αντίθετα, η σύνδεση με πολιτισμικό πλαίσιο, όπως φαίνεται στην E5, καταγράφει τον χαμηλότερο μέσο όρο του δείγματος (Μ.Ο. 2,67), με σημαντική διακύμανση και ελάχιστη βαθμολογία 1/5 από τον Μ8. Το γεγονός αυτό πιθανώς προκύπτει από τη φύση των δραστηριοτήτων – γραμματικές ασκήσεις στην πλειοψηφία τους– που επέλεξαν οι μαθητές, οι οποίες εστιάζουν στις δομές της ισπανικής γλώσσας και όχι στην πολιτισμική ή λογοτεχνική ερμηνεία του αποσπάσματος που χρησιμοποιήθηκε. Παρότι το απόσπασμα (μαγικός ρεαλισμός, Λατινική Αμερική) διαθέτει ισχυρή πολιτισμική διάσταση, αυτή δεν αναδείχθηκε στις ασκήσεις.

Η E8 (συνολική ικανοποίηση με Μ.Ο. 3,00) και η E7 (θα χρησιμοποιούσα ξανά αυτό το εργαλείο για να μελετήσω ισπανικά με Μ.Ο. 3,00) παραμένουν χαμηλά στην βαθμολογία, όμως με σημαντική διακύμανση. Η διακύμανση της E8 είναι ενδεικτική αφού ο Μ11 (Gemini) έδωσε 5/5, ενώ οι Μ7, Μ8 και Μ12 βαθμολόγησαν το σύστημα με 2/5. Η βαθμολογική αυτή απόκλιση δείχνει μια σαφή διαφοροποίηση μεταξύ των δύο γλωσσικών μοντέλων, αφού οι μαθητές που επέλεξαν με το Gemini, σε συνθήκες πλήρους λειτουργίας (Μ10, Μ11) δείχνουν περισσότερο ικανοποιημένοι από αυτούς που χρησιμοποίησαν το Llama (Μ7, Μ8, Μ9) και από τον Μ12 που αντιμετώπισε τεχνικά προβλήματα.

Μεγάλη διακύμανση παρουσιάζει και η E10 (Μ.Ο. 4,00), που αντιστοιχεί στην ενεργό συμμετοχή του μαθητή. Ο λόγος της μεγάλης διακύμανσης οφείλεται στο γεγονός ότι οι μαθητές Μ7, Μ8 και Μ11 βαθμολόγησαν το σύστημα με 5/5 και ο Μ12 με 2/5. Ο Μ12, προφανώς λόγω της απότομης διακοπής της συνεδρίας του, αντανακλά στην χαμηλή βαθμολογία του την εμπειρία αδυναμίας αλληλεπίδρασης.

5.4.3. Ανοιχτές Ερωτήσεις

Στις ερωτήσεις ανοιχτού τύπου, οι θετικές αναφορές στην E13 δείχνουν την άμεση σύνδεση των ασκήσεων με το απόσπασμα (*«οι ασκήσεις ήταν αποσπάσματα του κειμένου»*), τον Σωκρατικό διάλογο (*«με βοηθούσε να οδηγηθώ στη σωστή απάντηση ακόμη και όταν απάντησα με λάθη»*) και στην παραγωγή χρήσιμου υλικού για εργασία στο σπίτι.

Στην E14 (τι δεν λειτούργησε), οι αρνητικές αναφορές κινήθηκαν γύρω από τις γλωσσικές διαρροές (*«έβγαζε κάτι σαν κινέζικα»*, *«πολλές άγνωστες λέξεις σε άλλη*

γλώσσα»), την πρόωρη αποκάλυψη απαντήσεων και την τεχνική αποτυχία με χαρακτηριστική αναφορά του M12 στις κενές απαντήσεις που λάμβανε. Αξιοσημείωτη είναι και η παρατήρηση που δείχνει την μηχανιστική φύση του συστήματος (*«απαντάει μηχανήμα»*).

Στην E15 (διαφορά από τον καθηγητή), κυριαρχεί η ιδέα ότι ένας καθηγητής έχει δεξιότητες που η TN δεν διαθέτει. Χαρακτηριστικά, ένας μαθητής αναφέρει: *«δεν νομίζω ότι μπορεί να συγκριθεί»* δηλώνοντας έτσι τις έντονες διαφορές που εντόπισε μεταξύ εφαρμογής και εκπαιδευτικού.

Στην E16 (τι θα άλλαζαν), κυριαρχούν απαντήσεις που σχετίζονται με την μικρότερη έκταση των αποκρίσεων, με την απαλοιφή των markdown συμβόλων όπως #,* κ.λπ, με το μεγαλύτερο πλαίσιο διαλόγου και στη διεπαφή. Επίσης υπήρξε ένα αίτημα για ασκήσεις με στοιχεία παιχνιδοποίησης (*«να μοιάζουν πιο πολύ με παιχνίδι, για παράδειγμα όπως το Duolingo»*). Το τελευταίο αίτημα είναι πραγματικά ενδιαφέρον γιατί η παρούσα υλοποίηση είναι βασισμένη αποκλειστικά στον κειμενικό διάλογο. Ο μαθητής εδώ δίνει μια έξυπνη ιδέα για μελλοντική επέκταση της εφαρμογής.

5.4.4. Συνολική Εικόνα B1-B2

Η συνολική ανάλυση των δεδομένων για τα επίπεδα B1-B2 παρουσιάζει μια ενδιαφέρουσα εικόνα στο σύνολό της. Σύμφωνα με το αρχείο καταγραφής και τη γνώμη των μαθητών, το σύστημα λειτουργεί ικανοποιητικά στην παραγωγή εκπαιδευτικού υλικού (φυλλάδια, φύλλα εργασίας και ασκήσεις για το σπίτι) και παρουσιάζει στοιχεία Σωκρατικού διαλόγου σε γραμματικές δομές αυξημένης πολυπλοκότητας, όπως η υποτακτική. Το Gemini υπερτερεί του Llama αφού πετυχαίνει πιο συνεκτικό διάλογο και δείχνει να τηρεί τις οδηγίες του συστήματος προτροπών στις περισσότερες περιπτώσεις.

Ένα από τα πιο ενδιαφέροντα σημεία στην αξιολόγηση του συστήματος για τα επίπεδα B1-B2 είναι η χαμηλή βαθμολογία της E5 (M.O. 2,67), η χαμηλότερη του συνόλου του δείγματος σε αυτή την ερώτηση. Ο χαμηλός βαθμός πιθανώς φανερώνει ότι όταν οι μαθητές χρησιμοποιούν το σύστημα για εξάσκηση τυπικών γραμματικών φαινομένων, η λογοτεχνική και πολιτισμική διάσταση του κειμένου δεν αξιοποιείται στο μέγιστο. Το κενό αυτό ενδεχομένως να οφείλεται στο σύστημα προτροπών. Η αιτία πιθανώς βρίσκεται στην έλλειψη κάποιου κανόνα που θα εντάσσει ακόμα και τις γραμματικές ασκήσεις σε ένα λογοτεχνικό-ερμηνευτικό πλαίσιο.

Τέλος, τα τεχνικά προβλήματα ήταν πολύ πιο συχνά στα μεσαία επίπεδα. Η διακοπή μιας συνεδρίας λόγω 11 διαδοχικών κενών αποκρίσεων και η αντίστοιχη χαμηλή βαθμολογία στις E8, E9 και E10 του M12 υποδεικνύει ότι η μαθησιακή εμπειρία επηρεάζεται σημαντικά από την τεχνική αστάθεια των ψηφιακών συστημάτων.

5.5. Αποτελέσματα Αξιολόγησης των Εκπαιδευτικών

5.5.1. Ποσοτική Ανάλυση Ερωτηματολογίου

Δύο εκπαιδευτικοί ισπανικής γλώσσας (K1, K2) παρακολούθησαν τις συνεδρίες των μαθητών και συμπλήρωσαν το ερωτηματολόγιο αξιολόγησης που αποτελείται από 14 ερωτήσεις (βλέπε Παράρτημα III). Τα αποτελέσματα παρουσιάζονται στον Πίνακας 6.

Πίνακας 6: Απαντήσεις των καθηγητών K1 και K2 ανά τύπο ερώτησης: Ποσοτική αξιολόγηση βάσει 5βαθμης κλίμακας Likert (E1-E11, όπου 1: Διαφωνώ απόλυτα έως 5: Συμφωνώ απόλυτα) και συνολική/ποιοτική εκτίμηση της συνεδρίας (E12-E14, με διαβάθμιση από «Πολύ κακή» έως «Άριστη»). Με γαλάζια απόχρωση αποτυπώνονται οι ερωτήσεις που έλαβαν την ίδια βαθμολογία και από τους 2 εκπαιδευτικούς.

Ερώτηση	K1	K2
E1. Το σύστημα θέτει ερωτήσεις αντί να δίνει άμεσες απαντήσεις.	3	4
E2. Σε περίπτωση λάθους, δίνεται υπόδειξη αντί της σωστής απάντησης.	5	4
E3. Το επίπεδο καθοδήγησης είναι ανάλογο της δυσκολίας.	4	4
E4. Παρατηρείται σταδιακή μείωση υποστήριξης (fading) κατά τη συνεδρία.	2	2
E5. Το λεξιλόγιο και η γραμματική ταιριάζουν με το CEFR επίπεδο.	4	4
E6. Οι οδηγίες των ασκήσεων είναι πλήρεις και κατανοητές για τον μαθητή.	5	4
E7. Το σύστημα αναπροσαρμόζει τη δυσκολία βάσει των απαντήσεων του μαθητή.	3	3
E8. Η ανατροφοδότηση αναφέρεται στη συγκεκριμένη απάντηση του μαθητή.	5	5
E9. Στις απαντήσεις του συστήματος χρησιμοποιείται θετική ενίσχυση.	5	5
E10. Οι πληροφορίες που σχετίζονται με το κείμενο/συγγραφέα είναι ακριβείς.	5	4
E11. Οι ασκήσεις και τα παραδείγματα προέρχονται από το corpus.	4	4
E12. Πως κρίνω το σύστημα ως προς την παιδαγωγική αποτελεσματικότητα.	Μέτρια	Μέτρια

E13. Πώς κρίνω το σύστημα ως προς την ακρίβεια σε σχέση με το CEFR επίπεδο.	Καλή	Καλή
14. Πώς κρίνω το σύστημα ως προς την κινητροδότηση και την εμπλοκή του μαθητή.	Μέτρια	Καλή

Τα αποτελέσματα δείχνουν βαθμό σύγκλισης μεταξύ των δύο εκπαιδευτικών με ποσοστό απόλυτης συμφωνίας 64.3% (9 από τις 14 ερωτήσεις), ενώ οι αποκλίσεις στις υπόλοιπες ερωτήσεις δεν υπερέβαιναν τη μία μονάδα.

Ως προς τον Σωκρατικό διάλογο (E1–E4), οι δύο εκπαιδευτικοί αξιολογούν θετικά τη λειτουργία του συστήματος. Εξαίρεση αποτελεί η ερώτηση που σχετίζεται με την σταδιακή απόσυρση της υποστήριξης (fading), η οποία συγκεντρώνει τη χαμηλότερη βαθμολογία (2/5) στο σύνολο των ερωτήσεων. Η E4 αποτελεί και τη μοναδική περίπτωση απόλυτης συμφωνίας των K1 και K2 σε αρνητική αξιολόγηση. Δηλαδή, παρότι το σύστημα εφαρμόζει αποτελεσματικά στρατηγικές καθοδήγησης, σύμφωνα με τους εκπαιδευτικούς, δεν προσαρμόζεται στην σταδιακή απόσυρση καθώς ο μαθητής αποκτά μεγαλύτερη αυτονομία.

Στον δεύτερο άξονα, που αφορά το επίπεδο CEFR της εφαρμογής (E5–E7), οι εκπαιδευτικοί αξιολόγησαν θετικά την αντιστοιχία του λεξιλογίου και της γραμματικής με το επίπεδο των μαθητών καθώς και την πληρότητα και σαφήνεια των οδηγιών που έδωσε το σύστημα στις δραστηριότητες. Αντίθετα, η δυναμική προσαρμογή της δυσκολίας βάσει των αποκρίσεων του μαθητή έλαβε μέτρια βαθμολογία από τους K1 και K2, γεγονός που υποδηλώνει ότι η προσαρμοστικότητα του συστήματος και εδώ δεν λειτούργησε.

Η ποιότητα της ανατροφοδότησης του συστήματος στους μαθητές συγκέντρωσε τις υψηλότερες αξιολογήσεις του ερωτηματολογίου. Και οι δύο εκπαιδευτικοί συμφωνούν απόλυτα τόσο ως προς την εξατομικευμένη ανατροφοδότηση σε συγκεκριμένες απαντήσεις των μαθητών όσο και ως προς τη θετική ενίσχυση (οι E8 και E9 έλαβαν βαθμό 5/5 και από τους 2 εκπαιδευτικούς). Δηλαδή, η ανατροφοδότηση αποτελεί το ισχυρότερο παιδαγωγικό χαρακτηριστικό της εφαρμογής σύμφωνα με τους καθηγητές. Αντίστοιχα θετική είναι και η αξιολόγηση στις ερωτήσεις που αφορούν στην αξιοπιστία των παραγόμενων δραστηριοτήτων για το σπίτι (E10–E11). Οι K1 και K2 συμφωνούν ότι οι ασκήσεις και τα παραδείγματα που χρησιμοποίησε το σύστημα προέρχονται από

τα αποσπάσματα που χρησιμοποιήθηκαν, δηλαδή το σύστημα είναι σε γενικές γραμμές συνεπές ως προς το εκπαιδευτικό υλικό που παράγει.

Τέλος, στη συνολική αποτίμηση (E12–E14), οι δύο εκπαιδευτικοί αξιολογούν ομόφωνα την παιδαγωγική αποτελεσματικότητα του συστήματος ως «*Μέτρια*», ενώ ως προς τη γλωσσική ακρίβεια σε σχέση με τα επίπεδα του CEFR το χαρακτηρίζουν «*Καλό*». Η μοναδική διαφορά εντοπίστηκε στην κινητροδότηση και την εμπλοκή του μαθητή, όπου ο K1 χαρακτηρίζει το σύστημα «*Μέτριο*» και ο K2 «*Καλό*». Σε αυτό το σκέλος της έρευνας, παρατηρείται μια ουδέτερη στάση των εκπαιδευτικών, αφού κανένας από τους εκπαιδευτικούς δεν έδωσε ακραία θετικό («*Άριστη*») ή αρνητικό («*Πολύ κακή*») χαρακτηρισμό στο ερωτηματολόγιο.

5.5.2. Ποιοτική Ανάλυση — Ανοιχτές Ερωτήσεις

Στις ανοιχτές ερωτήσεις, ως προς την παιδαγωγική λογική και τον Σωκρατικό διάλογο, οι K1 και K2 εξέφρασαν αποκλίνουσες, και παράλληλα συμπληρωματικές, εκτιμήσεις για το σύστημα. Ο K1 αξιολογεί τον διάλογο ως «*μηχανιστικό στην πλειοψηφία των περιπτώσεων*», αναγνωρίζοντας ωστόσο εξαιρέσεις. Παράλληλα αποδίδει την συνέπεια του συστήματος στον σαφή παιδαγωγικό σχεδιασμό του («*ακολουθεί σαφείς οδηγίες στον τρόπο δημιουργίας ασκήσεων*»). Ο K2 θέτει άμεσα τη σύγκριση των δύο μοντέλων μεταξύ τους και αναγνωρίζει στοιχεία γνήσιου Σωκρατικού διαλόγου, ειδικά στο Gemini. Η τελευταία παρατήρηση βρίσκεται σε απόλυτη συμφωνία με τα δεδομένα του αρχείου καταγραφών, όπως αναφέρεται σε αρκετές περιπτώσεις και στις ενότητες 5.2.1., 5.3.1. και 5.4.1., αναμενόμενο αν θέσουμε την παράμετρο ότι οι καθηγητές είχαν πλήρη εικόνα των αρχείων καταγραφών πριν την συμπλήρωση των ερωτηματολογίων. Επίσης, και οι δύο εκπαιδευτικοί συμφωνούν ότι η εφαρμογή δουλεύει περισσότερο σαν εργαλείο παραγωγής ασκήσεων παρά ως γνήσιος παιδαγωγικός συνομιλητής.

Επίσης, και οι δύο εκπαιδευτικοί αξιολογούν ως «*Καλή*» την αντιστοίχιση του παραγόμενου από τα γλωσσικά μοντέλα υλικού ως προς το CEFR επίπεδο. Ο K1 επισημαίνει ότι «*στις περισσότερες περιπτώσεις οι ασκήσεις ταιριάζουν απόλυτα*», με μικρές αποκλίσεις κυρίως στα επίπεδα A1–A2. Ωστόσο, ο K2 παρατηρεί ένα κρίσιμο πρόβλημα στην αντιστοιχία επιπέδου λέγοντας ότι «*το σύστημα εμφάνισε πάρα πολλά γλωσσικά λάθη — άσχετες λέξεις εμφανίστηκαν στα κείμενα, γεγονός που δυσκόλευε τους μαθητές και τους αποσπούσε τη συγκέντρωση*». Η παρατήρηση αυτή επιβεβαιώνει τις

γλωσσικές διαρροές που εντοπίστηκαν και στα αρχεία καταγραφής και επαληθεύει ότι το φαινόμενο ήταν έντονα ορατό και από τον εξωτερικό παρατηρητή.

Στην ερώτηση για την εμπλοκή των μαθητών, ο Κ1 επισημαίνει ότι *«κάποιοι μαθητές εμφάνισαν σημάδια εκνευρισμού και ζήτησαν καθοδήγηση»* την ώρα της συνεδρίας, αξιολογώντας έτσι ως αδύναμη την ικανότητα του συστήματος να ανταποκριθεί στις ανάγκες των μαθητών κατά την διάρκεια της αλληλεπίδρασης. Ο Κ2 δήλωσε ότι στην πλειοψηφία τους οι μαθητές απλώς απαντούσαν διεκπεραιωτικά, όμως αναγνωρίζει εξαιρέσεις (*«σε ορισμένες περιπτώσεις ανέφεραν οι ίδιοι (οι μαθητές) ότι τους άρεσε η ανατροφοδότηση και η εξήγηση που πήραν»*). Η αντίφαση της μηχανικής αλληλεπίδρασης των μαθητών με το σύστημα σε συνδυασμό με την παρατήρηση ότι το σύστημα έχει αξία σε επίπεδο περιεχομένου, αντικατοπτρίζει τη διαφορά μεταξύ εμπλοκής και ικανοποίησης που φάνηκε και από τα ερωτηματολόγια μαθητών.

Τέλος, η πιο ξεκάθαρη διαφορά μεταξύ των δύο εκπαιδευτικών εντοπίζεται στις προοπτικές χρήσης. Ο Κ1 αποκλείει κατηγορηματικά τη χρήση του συστήματος στην τάξη (*«είναι κουραστικό ειδικά για τα μικρά επίπεδα»*), αλλά αναγνωρίζει την αξία του στην παραγωγή ασκήσεων και ειδικά στα *«θέματα έκθεσης»* και τις *«ασκήσεις για το σπίτι»*. Ο Κ2 υιοθετεί πιο ανοιχτή στάση δηλώνοντας ότι ενδεχομένως *«θα το χρησιμοποιούσα έπειτα από κάποιες αλλαγές»*. Επίσης, αναφέρει στα θετικά τις *«ωραίες ασκήσεις για το αντίστοιχο επίπεδο»* και τα *«ωραία παραδείγματα μέσα από τα κείμενα»* και στα αρνητικά του συστήματος, τη μεγάλη έκταση αποκρίσεων και τις γλωσσικές διαρροές. Κλείνοντας, και οι δύο εκπαιδευτικοί συμφωνούν απόλυτα ότι σε καμία περίπτωση δεν υποκαθιστά τον καθηγητή, δηλαδή το σύστημα έχει νόημα ως συμπληρωματικό εκπαιδευτικό μέσο και όχι ως υποκατάστατο του καθηγητή.

5.5.3. Συνολική Εικόνα από την Αξιολόγηση των Εκπαιδευτικών

Η αξιολόγηση των εκπαιδευτικών παρέχει και μια εξωτερική επικύρωση των ευρημάτων που προέκυψαν από τα αρχεία καταγραφής και τα ερωτηματολόγια των μαθητών. Συνοπτικά, άξια αναφοράς είναι η κοινή βαθμολογία 2/5 στην σταδιακή απόσυρση του συστήματος, η κοινή αξιολόγηση *«Μέτρια»* για την παιδαγωγική αποτελεσματικότητα του συστήματος και *«Καλή»* για την γλωσσική ακρίβεια που επιβεβαιώνει τα ευρήματα που ανέδειξε η μελέτη των αρχείων καταγραφής. Το σύστημα αποδίδει καλύτερα στη γλωσσική αντιστοίχιση επιπέδου παρά στην παιδαγωγική προσαρμογή στα ιδιαίτερα χαρακτηριστικά των μαθητών. Τέλος,

αμφότεροι συμφώνησαν ότι η εφαρμογή δεν αντικαθιστά τον καθηγητή αλλά έχει δυναμική ως συμπληρωματικό εργαλείο.

Κεφάλαιο 6: Συζήτηση των ευρημάτων

Με αυτό το κεφάλαιο επιχειρείται μια συζήτηση και μια κριτική σύνθεση των ευρημάτων όπως αυτά παρουσιάστηκαν στο κεφάλαιο 5, υπό το πρίσμα του θεωρητικού πλαισίου όπως αναπτύχθηκε στο κεφάλαιο 2 και -εν μέρει- στο κεφάλαιο 3 της παρούσας εργασίας.

6.1. Η TN ως Γνωστικό Εργαλείο

Το πρώτο ερευνητικό ερώτημα αφορά τον βαθμό στον οποίο η αλληλεπίδραση με το σύστημα υποστηρίζει, σύμφωνα με τους χρήστες-μαθητές, την κατανόηση του λογοτεχνικού κειμένου και την ανάπτυξη νέων γλωσσικών δεξιοτήτων. Τα αποτελέσματα της εργασίας δείχνουν ότι αυτό είναι εφικτό, δεν προκύπτει όμως αυτόματα από τη χρήση ενός μεγάλου γλωσσικού μοντέλου. Αντίθετα, προϋποθέτει σωστό παιδαγωγικό σχεδιασμό, κατάλληλες οδηγίες προς το γλωσσικό μοντέλο και δυνατότητα μνήμης, δηλαδή να αξιοποιεί τις προηγούμενες αλληλεπιδράσεις του με τον μαθητή κατά την συνεδρία.

Στην περίπτωσή μας, το Gemini κατάφερε να οδηγήσει τους μαθητές, μέσα από διαδοχικές ερωτήσεις και υποδείξεις, σε ερμηνευτικά συμπεράσματα που υπερβαίνουν τις απαιτήσεις του επιπέδου τους σε ορισμένες περιπτώσεις (π.χ. επίπεδα A1-A2). Από την άλλη μεριά, το Llama παρουσίασε μικρότερη παιδαγωγική συνέπεια, αφού σε αρκετές περιπτώσεις δεν διατήρησε τον Σωκρατικό διάλογο. Το εύρημα αυτό συμφωνεί με όσα περιγράφουν οι Gibson et al. (2023) για τη λειτουργία της TN στο micro-επίπεδο της μαθησιακής αλληλεπίδρασης. Επίσης συμφωνεί με αυτό που οι J. Liu et al. (2024) τεκμηρίωσαν εμπειρικά, ότι η μετάβαση από το απλό σχήμα «Ερώτηση–Απάντηση» σε έναν ουσιαστικό Σωκρατικό διάλογο δεν αποτελεί εγγενή ιδιότητα των μεγάλων γλωσσικών μοντέλων, αλλά απαιτεί συνειδητό παιδαγωγικό σχεδιασμό, είτε μέσω fine-tuning είτε μέσω μηχανικής προτροπών, στην περίπτωση μας.

Ωστόσο, η ίδια εικόνα δεν παρατηρήθηκε στα επίπεδα B1-B2 που η χαμηλή βαθμολογία στην αντίστοιχη ερώτηση, που αφορά τη σύνδεση με το ιστορικό και πολιτισμικό πλαίσιο, μας δείχνει ότι οι μαθητές οι οποίοι χρησιμοποίησαν το σύστημα κυρίως για την εξάσκηση γραμματικών φαινομένων δεν είχαν την ίδια άποψη. Εδώ η αλληλεπίδραση επικεντρώθηκε στη σωστή χρήση της γλώσσας και όχι στην ερμηνεία του λογοτεχνικού κειμένου. Το εύρημα αυτό δεν σχετίζεται με τις δυνατότητες των

γλωσσικών μοντέλων, αλλά οφείλεται στον παιδαγωγικό σχεδιασμό του συστήματος προτροπών. Οι οδηγίες που δόθηκαν δεν καθοδηγούσαν το μοντέλο με σαφήνεια να συνδέει τις γραμματικές ασκήσεις με τη λογοτεχνική και πολιτισμική διάσταση του κειμένου. Γι' αυτό, το σύστημα ακολουθούσε το άμεσο αίτημα του μαθητή, χωρίς να αξιοποιεί την ευκαιρία για βαθύτερη ερμηνευτική συζήτηση, κενό που εντοπίζεται στον παιδαγωγικό σχεδιασμό του συστήματος προτροπών. Συνολικά, η παρατήρηση αυτή συμφωνεί με τους Gibson et al. (2023), οι οποίοι υποστηρίζουν ότι η TN λειτουργεί αποτελεσματικά ως μαθησιακό εργαλείο όταν ο παιδαγωγικός σχεδιασμός καθορίζει με σαφήνεια τους μαθησιακούς στόχους που καλείται να υπηρετήσει.

Συνεπώς, ως προς το EE1, η απάντηση είναι υπό όρους. Το προτεινόμενο σύστημα συμβάλει αποτελεσματικά στην κατανόηση λογοτεχνικού κειμένου και την γλωσσική ανάπτυξη του μαθητή, όταν η αλληλεπίδραση ευθυγραμμίζεται με τους σχεδιαστικούς και τους παιδαγωγικούς στόχους που θέτονται. Οποιαδήποτε απόκλιση αυτής της ευθυγράμμισης, οδηγεί σε χαμηλότερη παιδαγωγική αξία, υπογραμμίζοντας πως η σκαλωσιά εξαρτάται και από τον αρχιτεκτονικό σχεδιασμό του συστήματος προτροπών, την μνήμη του γλωσσικού μοντέλου και από τις ικανότητες του ίδιου του γλωσσικού μοντέλου.

6.2. Παιδαγωγική Αποτελεσματικότητα και Χρηστικότητα

Το δεύτερο ερευνητικό ερώτημα αφορά στην αξιολόγηση της παιδαγωγικής αποτελεσματικότητας και της χρηστικότητας της εφαρμογής από μαθητές και εκπαιδευτικούς. Εδώ εμφανίζεται μια δομική ασυμμετρία που η βιβλιογραφία βοηθά να ερμηνευθεί. Παράλληλα, η σύγκριση των δύο γλωσσικών μοντέλων μας δείχνει ότι η αρχιτεκτονική επιλογή αποτελεί κρίσιμη παιδαγωγική απόφαση.

Η μαθησιακή εμπλοκή ήταν ιδιαίτερα υψηλή και στα τρία επίπεδα γλωσσομάθειας CEFR, γεγονός που μας δείχνει ότι οι μαθητές συμμετείχαν ενεργά στη μαθησιακή διαδικασία, ανεξάρτητα από το επίπεδό της γλώσσας στο οποίο βρίσκονται. Το εύρημα αυτό ευθυγραμμίζεται με τον ισχυρισμό του El-Sabagh (2021), ο οποίος υποστηρίζει ότι τα προσαρμοστικά μαθησιακά περιβάλλοντα ευνοούν την εμπλοκή των μαθητών, επειδή τους επιτρέπουν να μαθαίνουν με τον δικό τους ρυθμό. Ιδιαίτερο ενδιαφέρον παρουσιάζει το γεγονός ότι ορισμένοι μαθητές δήλωσαν πως ένιωθαν άνετα να κάνουν ερωτήσεις που πιθανόν θα δίσταζαν να θέσουν μέσα στην τάξη. Το αποτέλεσμα αυτό συμφωνεί με την Fadieieva (2023), σύμφωνα με την οποία η αίσθηση ψυχολογικής

ασφάλειας αποτελεί βασική προϋπόθεση για την ενεργή συμμετοχή στη μάθηση. Σε αντίθεση με έναν άνθρωπο (συμμαθητή ή εκπαιδευτικό), το σύστημα δεν κρίνει αρνητικά και δεν συγκρίνει τον χρήστη, γεγονός που φαίνεται να μειώνει το άγχος και να ενθαρρύνει τη συμμετοχή.

Αντίθετα με την μαθησιακή εμπλοκή, η συνολική ικανοποίηση από το σύστημα και η πρόθεση επαναχρησιμοποίησης του ήταν πιο συγκρατημένες, ειδικά στα επίπεδα B1-B2. Το εύρημα αυτό είναι λίγο αντιφατικό συγκριτικά με την μαθησιακή εμπλοκή όπως αποτυπώθηκε από τους μαθητές. Πώς ένας μαθητής μπορεί να συμμετέχει ενεργά στη μαθησιακή διαδικασία χωρίς απαραίτητα να είναι απόλυτα ικανοποιημένος από την εμπειρία; Η αντίφαση θα μπορούσε να ερμηνευτεί στη βάση του μοντέλου ARCS του Keller (1987), σύμφωνα με το οποίο η εμπλοκή και η ικανοποίηση δεν ταυτίζονται απαραίτητα. Στη δική μας περίπτωση, η χαμηλότερη ικανοποίηση φαίνεται να σχετίζεται κυρίως με τεχνικά προβλήματα, όπως κενές αποκρίσεις ή ελλείψεις και προβλήματα στο σχεδιασμό της διεπαφής, οι οποίες επηρέασαν τη συνολική εμπειρία των μαθητών. Το συμπέρασμα αυτό έρχεται να ενισχύσει η σύγκριση των δύο μοντέλων. Οι μαθητές που δούλεψαν με το Gemini, στις περιπτώσεις που δεν παρουσίασε τεχνικά προβλήματα, αξιολόγησαν θετικότερα την εμπειρία τους. Συνεπώς, η παιδαγωγική αξία δεν εξαρτάται μόνο από την ποιότητα καθοδήγησης αλλά και από την σταθερότητα τέτοιων συστημάτων. Την ίδια εικόνα επιβεβαίωσαν στην αξιολόγησή τους και οι δύο εκπαιδευτικοί. Εκεί που οι μαθητές σημείωσαν υψηλή εμπλοκή, οι εκπαιδευτικοί είδαν καλή ανατροφοδότηση και θετική ενίσχυση και εκεί που οι μαθητές ένιωσαν ότι το σύστημα δεν ανταποκρίθηκε στις προσδοκίες τους, οι εκπαιδευτικοί το επιβεβαίωσαν με τη χαμηλή βαθμολογία που έδωσαν στο σύστημα.

Ένα ακόμη ζήτημα που χρήζει συζήτησης και επηρεάζει άμεσα την παιδαγωγική εμπειρία είναι η διαφορά στην αξιοπιστία των δύο μοντέλων. Το Llama υποστηρίζει επίσημα οκτώ γλώσσες (Αγγλικά, Γερμανικά, Γαλλικά, Ιταλικά, Πορτογαλικά, Ισπανικά, Χίντι και Ταϊλανδέζικα) στις οποίες δεν περιλαμβάνονται τα Ελληνικά²². Το Gemini βασίζεται σε αρχιτεκτονική sparse Mixture-of-Experts με πολυγλωσσική κάλυψη άνω των 400 γλωσσών²³. Αυτό έχει άμεσες παιδαγωγικές συνέπειες, γιατί το σύστημα χρησιμοποιεί τρεις γλώσσες ταυτόχρονα (Αγγλικά για τις οδηγίες, Ελληνικά για την έξοδο, Ισπανικά για το κείμενο και τις ασκήσεις). Οι Huang et al. (2023) έδειξαν

²² <https://huggingface.co/meta-llama/Llama-3.3-70B-Instruct>

²³ https://storage.googleapis.com/deepmind-media/gemini/gemini_v2_5_report.pdf

ότι τα γλωσσικά μοντέλα δεν επεξεργάζονται εξίσου αποτελεσματικά όλες τις γλώσσες, και οι Shi et al. (2022) ότι η παιδαγωγική λογική αποδίδεται καλύτερα όταν οι οδηγίες είναι στα Αγγλικά. Τα ευρήματά μας επιβεβαιώνουν αυτές τις παρατηρήσεις αφού οι γλωσσικές διαρροές του Llama δεν είναι τυχαία σφάλματα αλλά συνέπεια της γλωσσικής εκπαίδευσής, φαινόμενο που ο Z. Liu (2023) περιγράφει ως αλγοριθμική μεροληψία. Το συμπέρασμα είναι ότι η επιλογή γλωσσικού μοντέλου σε ένα πολύγλωσσο εκπαιδευτικό περιβάλλον δεν είναι αμιγώς τεχνικής φύσης απόφαση αλλά έντονα παιδαγωγική.

6.3. Σωκρατικός Διάλογος, Σκαλωσιά και Σταδιακή Απόσυρση

Το τρίτο ερευνητικό ερώτημα αφορά τον βαθμό στον οποίο το σύστημα εφαρμόζει στην πράξη τις αρχές του Σωκρατικού διαλόγου, της σκαλωσιάς και της σταδιακής απόσυρσης. Το Gemini εφάρμοσε με συνέπεια την σκαλωσιά κατευθύνοντας τους μαθητές χωρίς να τους αποκαλύπτει άμεσα την απάντηση στα ερωτήματα που έθετε. Δηλαδή λειτούργησε ως «ικανότερος άλλος» (βλέπε Ενότητα 2.2.2.). Αυτή η συμπεριφορά αντιστοιχεί σε αυτό που ο Margolis (2020) αναγνωρίζει ως την πλέον αποτελεσματική εκδοχή της ZEA της θεωρίας του Vygotsky στη διδακτική πράξη. Το Llama παρουσίασε το ακριβώς αντίθετο μοτίβο, αφού αποκάλυπτε συστηματικά τις απαντήσεις στα ερωτήματα που έθετε, είτε μέσα στην εκφώνηση είτε έπειτα από προτροπή του μαθητή για βοήθεια. Το φαινόμενο αυτό συνδέεται με αυτό που η βιβλιογραφία αναγνωρίζει ως helpfulness bias (Z. Liu, 2023; Mollick & Mollick, 2023), δηλαδή την τάση των μεγάλων γλωσσικών μοντέλων να παρέχουν την πληρέστερη δυνατή απάντηση ακόμα και όταν αυτό αντιβαίνει στις παιδαγωγικές οδηγίες που έχουν λάβει.

Η αντίθεση αυτή είναι ιδιαίτερα διαφωτιστική στο επίπεδο C1-C2. Εδώ, η σκαλωσιά είχε τις περισσότερες πιθανότητες να καταρρεύσει, καθώς η απουσία μιας μοναδικής ορθής απάντησης, λόγω της σκόπιμης αοριστίας του κειμένου, στερεί από το μοντέλο το σαφές σημείο αναφοράς πάνω στο οποίο βασίζεται η καθοδήγηση. Παρ' όλα αυτά, όπως τεκμηριώθηκε στην ενότητα 5.3.1, το Gemini διατήρησε ακέραιο τον Σωκρατικό διάλογο ακόμη και υπό την άμεση πίεση του μαθητή για έτοιμες απαντήσεις, ενώ το Llama υπέκυψε για ακόμη μια φορά στο helpfulness bias. Το γεγονός ότι το ίδιο απαιτητικό κείμενο, οδήγησε σε διαμετρικά αντίθετα αποτελέσματα ανάλογα με το μοντέλο, ενισχύει το κεντρικό συμπέρασμα. Δηλαδή, η επιτυχία της σκαλωσιάς δεν

είναι συνάρτηση της δυσκολίας του κειμένου, αλλά της παιδαγωγικής συνέπειας του γλωσσικού μοντέλου που χρησιμοποιείται κάθε φορά.

Ως προς τη σταδιακή απόσυρση, το σύστημα παρουσίασε έντονα το πρόβλημα και για τα δύο γλωσσικά μοντέλα, σύμφωνα και με την ομόφωνη χαμηλή βαθμολογία που έλαβε και από τους δύο εκπαιδευτικούς. Στο σύστημα προτροπών υπάρχει ρητή οδηγία να μειώνει σταδιακά την καθοδήγηση καθώς ο μαθητής αποκτά αυτονομία. Η οδηγία αυτή δεν εφαρμόστηκε γιατί ενδεχομένως στον σχεδιασμό του συστήματος δεν έγινε πρόβλεψη για μόνιμη μνήμη. Η Verenikina (2003) ισχυρίζεται ότι η σταδιακή απόσυρση απαιτεί τη διαχρονική γνώση της πορείας του μαθητή. Η τεχνική αιτία της αδυναμίας αυτής εντοπίζεται στην ίδια την αρχιτεκτονική του συστήματος, η οποία λειτουργεί με λογική stateless: κάθε συνεδρία στηρίζεται σε ανεξάρτητες κλήσεις προς τα API των μοντέλων, χωρίς αποθήκευση και ανάκτηση του ιστορικού του μαθητή σε βάση δεδομένων. Ο περιορισμός αυτός δεν είναι μοναδικός της παρούσας υλοποίησης. Αποτελεί γνωστό δομικό πρόβλημα των εκπαιδευτικών εφαρμογών που βασίζονται σε μεγάλα γλωσσικά μοντέλα χωρίς μνήμη, εμπόδιο που έχει παρατηρηθεί στην προσαρμοστική μάθηση (Zhang et al., 2025; Gibson et al., 2023). Επίσης, είναι δίκαιο στο σημείο αυτό να αναφέρουμε ότι η διαδικασία που ακολουθήθηκε κατά την αξιολόγηση δεν ενδείκνυται για ασφαλή συμπεράσματα, αφού ο κάθε μαθητής είχε μια μοναδική συνεδρία με το σύστημα.

Τέλος, ένα λογικό ερώτημα είναι κατά πόσον η ερμηνευτική ασάφεια των λογοτεχνικών κειμένων ευνόησε την εμφάνιση ψευδαισθήσεων. Η απάντηση δεν είναι μονοσήμαντη. Χαρακτηριστικό παράδειγμα αποτελεί το απόσπασμα «*Casa tomada*» του Cortázar, ένα ανοιχτό σε πολλαπλές ερμηνείες κείμενο, σύμφωνα με τον ίδιο τον συγγραφέα (βλέπε ενότητα 4.3.2.). Εκεί, οι αναφορές των μοντέλων σε λέξεις και χωρία του διηγήματος παρέμειναν ακριβείς και η πολυσημία του κειμένου δεν οδήγησε σε επινόηση ανύπαρκτων στοιχείων. Αντίθετα, φάνηκε να εντείνει το helpfulness bias του Llama, το οποίο επικύρωνε κάθε υπόθεση του μαθητή. Σε ένα κείμενο χωρίς μοναδική ορθή ερμηνεία, η αστοχία αυτή είναι παιδαγωγικά εξίσου επιζήμια με τις ψευδαισθήσεις, καθώς στερεί από τον μαθητή τη γνωστική σύγκρουση που απαιτεί ένα μη-δομημένο πρόβλημα (Loyens & Gijbels, 2008). Το Gemini, πάνω στο ίδιο κείμενο, δεν εμφάνισε αντίστοιχη συμπεριφορά.

Η απουσία επινόησης κειμενικών γεγονότων δεν σημαίνει, ωστόσο, ότι οι συνεδρίες με το Llama ήταν απαλλαγμένες από ψευδαισθήσεις. Σε γενικές γραμμές εντοπίστηκαν

δύο τύποι ψευδαισθήσεων. Αποκρίσεις που αντιφάσκουν με τις οδηγίες του ίδιου του συστήματος και παραγόμενο εκπαιδευτικό υλικό εντελώς ασύνδετο με το λογοτεχνικό κείμενο που ορίστηκε ως βάση της συνεδρίας, ειδικά στα επίπεδα C1-C2. Το πρόβλημα με αυτού του είδους τις ψευδαισθήσεις είναι έντονο σε ένα εκπαιδευτικό πλαίσιο χρήσης. Ο μαθητής, εξ ορισμού, δεν διαθέτει ακόμη την επάρκεια που θα του επέτρεπε να αναγνωρίσει το σφάλμα και ενδέχεται να εσωτερικεύσει και να αφομοιώσει την εσφαλμένη πληροφορία ως έγκυρη. Το εύρημα αυτό επιβεβαιώνει τα δύο είδη hallucination που περιγράφουν οι Ji et al. (2023) και αποτελεί πρόσθετο επιχείρημα υπέρ της προσεκτικής επιλογής μοντέλου σε εκπαιδευτικά περιβάλλοντα.

Κεφάλαιο 7: Συμπεράσματα και Μελλοντικές Επεκτάσεις

Στην παρούσα εργασία σχεδιάσε, σχεδιάστηκε, υλοποιήθηκε και αξιολογήθηκε παιδαγωγικά ένα σύστημα TN για τη διδασκαλία της ισπανικής γλώσσας μέσα από λογοτεχνικά κείμενα. Από τη συνδυαστική ανάλυση των αρχείων καταγραφής, των ερωτηματολογίων και της αξιολόγησης των εκπαιδευτικών, προκύπτει ότι αυτού του είδους τα συστήματα έχουν παιδαγωγική αξία αλλά βρίσκονται ακόμη σε προοίμιο στάδιο για να εφαρμοστούν στην πράξη. Η παιδαγωγική αξία ενός τέτοιου συστήματος δεν προκύπτει ως εγγενής ιδιότητα των μεγάλων γλωσσικών, αλλά ως προϊόν του παιδαγωγικού σχεδιασμού του συστήματος προτροπών, της σταθερότητας της υλοποίησης και της επιλογής του κατάλληλου μοντέλου.

Ως προς το EE1, το σύστημα υποστηρίζει αποτελεσματικά την κατανόηση του κειμένου και τη γλωσσική ανάπτυξη των μαθητών. Αυτό όμως ισχύει υπό μία προϋπόθεση. Η αλληλεπίδραση πρέπει να ευθυγραμμίζεται με τους παιδαγωγικούς στόχους. Όταν ο μαθητής στρέφει τη συνεδρία σε καθαρά γραμματική εξάσκηση, η λογοτεχνική και πολιτισμική διάσταση μένει αναξιοποίητη. Ως προς το EE2, αναδείχθηκε μια δομική ασυμμετρία. Η μαθησιακή εμπλοκή ήταν σταθερά υψηλή σε όλα τα επίπεδα αλλά η συνολική ικανοποίηση παρέμεινε πιο συγκρατημένη. Αυτό οφείλεται κυρίως στα τεχνικά προβλήματα. Ως προς το EE3, το σύστημα υλοποιεί με συνέπεια τη σκαλωσιά και τον Σωκρατικό διάλογο κάτι που ισχύει ιδίως για το γλωσσικό μοντέλο Gemini. Όμως το σύστημα αποτυγχάνει συστηματικά στη σταδιακή απόσυρση (fading). Ο περιορισμός αυτός συνδέεται άμεσα με την απουσία διαχρονικής μνήμης. Το ίδιο ισχύει και για τις ψευδαισθήσεις και τη μεροληψία. Τα δύο μοντέλα δεν συμπεριφέρθηκαν με τον ίδιο τρόπο αφού το Gemini παρέμεινε αξιόπιστο, χωρίς γλωσσικές διαρροές ενώ το Llama εμφάνισε μεροληψία στη γλώσσα και παρήγαγε υλικό αποκομμένο από το κείμενο.

Ως προς την συμβολή της, η εργασίας παρέχει μια λειτουργική αρχιτεκτονική που μεταφράζει κονστрукτιβιστικές αρχές σε μηχανική προτροπών και παράλληλα τεκμηριώνει εμπειρικά ότι η επιλογή γλωσσικού μοντέλου και της γλώσσας αλληλεπίδρασης αποτελούν παιδαγωγικές —και όχι απλώς τεχνικές— αποφάσεις.

Οι βασικοί περιορισμοί σχετίζονται το μικρό δείγμα (12 μαθητές, 2 εκπαιδευτικοί, πιλοτική μελέτη), η εξάρτηση από εξωτερικά API με τους περιορισμούς της δωρεάν έκδοσης και η απουσία μόνιμης μνήμης.

Κλείνοντας, υπάρχουν αρκετά περιθώρια για μελλοντική βελτίωση του συστήματος. Πρώτα από όλα, το σύστημα χρειάζεται μόνιμη μνήμη για κάθε μαθητή, ώστε να θυμάται την πορεία του και να μπορεί να αποσύρει σταδιακά την υποστήριξη. Επιπλέον, οι οδηγίες προς το μοντέλο θα μπορούσαν να εμπλουτιστούν με σαφείς κανόνες, που θα συνδέουν τις δραστηριότητες με το λογοτεχνικό και πολιτισμικό νόημα του κειμένου ανεξάρτητα από την ίδια την φύση τους (γραμματική, συντακτικό κ.λπ.). Χρήσιμη θα ήταν και η προσθήκη στοιχείων παιχνιδοποίησης, ειδικά στα πιο χαμηλά επίπεδα. Τέλος, το σύστημα αξίζει να δοκιμαστεί σε μεγαλύτερη κλίμακα και με περισσότερα μοντέλα.

Βιβλιογραφία

- Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). *IEEE access*, 6, 52138-52160.
- Almulla, M. A. (2023). Constructivism learning theory: A paradigm for students' critical thinking, creativity, and problem solving to affect academic performance in higher education. *Cogent Education*, 10(1), 2172929.
- Bayounes, W., Saâdi, I. B., & Kinshuk. (2022). Adaptive learning: toward an intentional model for learning process guidance based on learner's motivation. *Smart Learning Environments*, 9(1), 33.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77-101.
- Bell-Villada, G. H. (2010). *García Márquez: The man and his work*. Univ of North Carolina Press.
- Chapelle, C., & Chapelle, C. A. (2001). *Computer applications in second language acquisition*. Cambridge university press.
- Cohen, L., Manion, L., & Morrison, K. (2002). *Research methods in education*. routledge.
- Cole, M., & SCRIBNER, S. (1978). *Vygotsky, Lev S.(1978): Mind in Society. The Development of Higher Psychological Processes*.
- Creswell, J. W., & Clark, V. P. (2007). Mixed methods research. *Thousand Oaks, CA*.
- Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: the state of the field. *International journal of educational technology in higher education*, 20(1), 1-22.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 13(3), 319-340.
- Denzin, N. K. (2012). Triangulation 2.0. *Journal of mixed methods research*, 6(2), 80-88.

- El-Sabagh, H. A. (2021). Adaptive e-learning environment based on learning styles and its impact on development students' engagement. *International Journal of Educational Technology in Higher Education*, 18(1), 53.
- Fadeev, A. (2019). Vygotsky's theory of mediation in digital learning environment: Actuality and practice. *Punctum*, 5(1), 24–44.
- Fadieieva, L. O. (2023). Adaptive learning: a cluster-based literature review (2011-2022). *Educational Technology Quarterly*, 2023(3), 319–366.
- Fahad Mon, B., Wasfi, A., Hayajneh, M., Slim, A., & Abu Ali, N. (2023, September). Reinforcement learning in education: A literature review. In *Informatics* (Vol. 10, No. 3, p. 74). MDPI.
- Faris, W. B. (2004). *Ordinary enchantments: Magical realism and the remystification of narrative*. Vanderbilt university press.
- Fielding, R. T. (2000). *Architectural styles and the design of network-based software architectures*. University of California, Irvine.
- Fitria, T. N. (2021, December). Artificial intelligence (AI) in education: Using AI tools for teaching and learning process. In *Prosiding seminar nasional & call for paper STIE AAS* (pp. 134-147).
- Gibson, D., Kovanovic, V., Ifenthaler, D., Dexter, S., & Feng, S. (2023). Learning theories for artificial intelligence promoting learning processes. *British Journal of Educational Technology*, 54(5), 1125–1146.
- Guo, H., Yi, W., & Liu, K. (2024). Enhancing constructivist learning: the role of generative AI in personalised learning experiences. *Proceedings of the 26th International Conference on Enterprise Information Systems*, 767–770.
- Håkansson, A., & Phillips-Wren, G. (2024). Generative AI and large language models- benefits, drawbacks, future and recommendations. *Procedia Computer Science*, 246, 5458-5468.
- Havard, B., East, M. L., Prayaga, L., & Whiteside, A. (2016). Adaptable learning theory framework for technology-enhanced learning. In *Handbook of research on applied learning theory and design in modern education* (pp. 632–654). IGI Global Scientific Publishing.

- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Huang, H., Tang, T., Zhang, D., Zhao, W. X., Song, T., Xia, Y., & Wei, F. (2023, December). Not all languages are created equal in llms: Improving multilingual capability by cross-lingual-thought prompting. In *Findings of the Association for Computational Linguistics: EMNLP 2023* (pp. 12365-12394).
- Hwang, G. J. (2014). Definition, framework and research issues of smart learning environments-a context-aware ubiquitous learning perspective. *Smart Learning Environments*, 1(1), 4.
- Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., ... & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM computing surveys*, 55(12), 1-38.
- Jonassen, D. H. (1997). Instructional design models for well-structured and III-structured problem-solving learning outcomes. *Educational Technology Research and Development*, 45(1), 65–94.
- Keller, J. M. (1987). Development and use of the ARCS model of instructional design. *Journal of instructional development*, 10(3), 2-10.
- Liu, J., Huang, Z., Xiao, T., Sha, J., Wu, J., Liu, Q., ... & Chen, E. (2024). SocraticLM: Exploring socratic personalized teaching with large language models. *Advances in Neural Information Processing Systems*, 37, 85693-85721.
- Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM computing surveys*, 55(9), 1-35.
- Liu, Z. (2023). Cultural bias in large language models: A comprehensive analysis and mitigation strategies. *Journal of Transcultural Communication*, 3(2), 224–244.
- Loyens, S. M., & Gijbels, D. (2008). Understanding the effects of constructivist learning environments: Introducing a multi-directional approach. *Instructional Science*, 36(5), 351–357.
- Margolis, A. A. (2020). Zone of Proximal Development, Scaffolding and Teaching Practice. *Cultural-Historical Psychology*, 16(3).

- Mollick, E., & Mollick, L. (2023). Assigning AI: Seven approaches for students, with prompts. *arXiv preprint arXiv:2306.10052*.
- Paul, R., & Elder, L. (2008). Critical thinking: The art of Socratic questioning, part III. *Journal of developmental education*, 31(3), 34-35.
- Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training.
- Rasul, Tareq, Sumesh Nair, Diane Kalendra, Mulyadi Robin, Fernando de Oliveira Santini, Wagner Ladeira, Mingwei Sun, Ingrid Day, Raouf Ahmad Rather, and Liz Heathcote. "The role of ChatGPT in higher education: Benefits, challenges, and future research directions." *Journal of Applied Learning & Teaching* 6, no. 1 (2023): 41-56.
- Saarsar, P. (2018). Exploring the constructivist approach in education: Theory, practice, and implications. *International Journal of Research and Analytical Reviews*, 5(2), 716–724.
- Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Design and evaluation of a short version of the user experience questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence*, 4(6), 103-108.
- Shi, F., Suzgun, M., Freitag, M., Wang, X., Srivats, S., Vosoughi, S., ... & Wei, J. (2022). Language models are multilingual chain-of-thought reasoners. *arXiv preprint arXiv:2210.03057*.
- Standish, P. (2001). *Understanding Julio Cortázar*. Univ of South Carolina Press.
- Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research, and Evaluation*, 9(1).
- Stoeckel, M. R. (2020). Literature Review of Constructivism in Online Science Courses. *ArXiv Preprint ArXiv:2007.07745*.
- Thakur, M. K. (2024). A Critical Review of Elemental Odes by Pablo Neruda. *Creative Saplings*, 3(02), 19-28.
- Tilkov, S., & Vinoski, S. (2010). Node.js: Using JavaScript to build high-performance network programs. *IEEE Internet Computing*, 14(6), 80-83.
- Verenikina, I. (2003). *Understanding scaffolding and the ZPD in educational research*.

- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., ... & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35, 24824-24837.
- White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., ... & Schmidt, D. C. (2023). A prompt pattern catalog to enhance prompt engineering with chatgpt. *arXiv preprint arXiv:2302.11382*.
- Yanthi, E. Y., Elmustian, E., & Charlina, C. (2024). The Effect of ChatGPT-based Constructivist Learning Model on Observation Report Text Writing Skills. *Journal of Languages and Language Teaching*, 12(4), 1952–1964.
- Yilmaz, K. (2008). Constructivism: Its theoretical underpinnings, variations, and implications for classroom instruction. *Educational Horizons*, 86(3), 161–172.
- Zhang, Z., Dai, Q., Bo, X., Ma, C., Li, R., Chen, X., ... & Wen, J. R. (2025). A survey on the memory mechanism of large language model-based agents. *ACM Transactions on Information Systems*, 43(6), 1-47.
- Zhao, L., & Wang, H. (2019, August). Research on adaptive learning system based on three core modules. In *2019 10th International Conference on Information Technology in Medicine and Education (ITME)* (pp. 447-452). IEEE.

Παράρτημα I

Ερωτήσεις Γενικού Τύπου

Ερωτήσεις που σχετίζονται με το προφίλ του χρήστη.

fgerodim@gmail.com [Εναλλαγή λογαριασμού](#)

Δεν κοινοποιήθηκε

* Υποδεικνύει απαιτούμενη ερώτηση

1. Ηλικία *

☐ 12-18

☐ 18-25

☐ 25-45

☐ πάνω από 45

2. Επίπεδο σπουδών *

☐ Δευτεροβάθμια Εκπαίδευση

☐ Τριτοβάθμια Εκπαίδευση

☐ Μεταπτυχιακό/Διδακτορικό

3. Επίπεδο στο οποίο βρίσκεστε στην ισπανική γλώσσα. *

☐ A1 - A2

☐ B1 - B2

☐ Γ1 - Γ2

4. Για ποιο λόγο επιλέξατε την ισπανική γλώσσα; *

☐ Ενασχόληση με την ισπανόφωνη λογοτεχνία και τον πολιτισμό

☐ Ταξίδι ή διαμονή σε ισπανόφωνη χώρα

☐ Δεύτερο ξενόγλωσσο τίτλο σπουδών για μοριοδότηση

☐ Άλλος λόγος

5. Γνωρίζετε τουλάχιστον άλλη μια ξένη γλώσσα πέρα από την ισπανική; *

☐ Ναι

☐ Όχι

6. Σε περίπτωση που **απαντήσατε Ναι** στην προηγούμενη ερώτηση, ποιο είναι το ανώτατο επίπεδό σας;

☐ A1 - A2

☐ B1 - B2

☐ Γ1 - Γ2

7. Πόσο συχνά έχετε χρησιμοποιήσει στο παρελθόν AI εργαλεία για μελέτη ή εκμάθηση (π.χ. ChatGPT, Duolingo, Google Translate); *

	1	2	3	4	
Ποτέ	☐	☐	☐	☐	Πολύ συχνά

8. Πιστεύετε ότι οι νέες τεχνολογίες όπως η Τεχνητή Νοημοσύνη μπορεί να συνεισφέρει στην εκμάθηση μιας ξένης γλώσσας; *

	1	2	3	4	5	
Διαφωνώ Απόλυτα	☐	☐	☐	☐	☐	Συμφωνώ Απόλυτα

Υποβολή

Εκκαθάριση φόρμας

Αυτό το περιεχόμενο δεν έχει δημιουργηθεί και δεν έχει εγκριθεί από την Google. · [Κάταχος φόρμας](#)
[επικοινωνίας](#) · [Όροι παροχής υπηρεσιών](#) · [Πολιτική απορρήτου](#)

Αυτή η φόρμα φαίνεται ύποπτη. [Αναφορά](#)

Google Φόρμες

Παράρτημα II

Ερωτηματολόγιο αξιολόγησης μαθητή

Μέρος 1ο: Χρησιμότητα και ευκολία χρήσης

fgerodim@gmail.com [Εγγραφή λογαριασμού](#)

Δεν κοινοποιήθηκε

* Υποδεικνύει απαιτούμενη ερώτηση

1. Ο ψηφιακός βοηθός ήταν απλός και κατανοητός στην χρήση του. *

1

2

3

4

5

Διαφωνώ απόλυτα.

☐

☐

☐

☐

☐

Συμφωνώ απόλυτα.

2. Η ταχύτητα απόκρισης του συστήματος ήταν ικανοποιητική. *

1

2

3

4

5

Διαφωνώ απόλυτα.

☐

☐

☐

☐

☐

Συμφωνώ απόλυτα.

3. Οι απαντήσεις που πήρα αντιστοιχούσαν στο επίπεδό μου. *

1

2

3

4

5

Διαφωνώ απόλυτα.

☐

☐

☐

☐

☐

Συμφωνώ απόλυτα.

4. Το σύστημα δεν μου έδινε έτοιμες απαντήσεις — με καθοδηγούσε να τις βρω μόνος μου. *

1

2

3

4

5

Διαφωνώ απόλυτα.

☐

☐

☐

☐

☐

Συμφωνώ απόλυτα.

5. Το σύστημα με βοήθησε να κατανοήσω καλύτερα το κείμενο και να συνδέσω το κείμενο με το ιστορικό και πολιτισμικό του πλαίσιο. *

1 2 3 4 5
Διαφωνώ απόλυτα. ☐ ☐ ☐ ☐ ☐ Συμφωνώ απόλυτα.

6. Ο ψηφιακός βοηθός μου κίνησε το ενδιαφέρον για να ασχοληθώ με το κείμενο. *

1 2 3 4 5
Διαφωνώ απόλυτα. ☐ ☐ ☐ ☐ ☐ Συμφωνώ απόλυτα.

7. Θα χρησιμοποιούσα ξανά αυτό το εργαλείο για να μελετήσω ισπανικά λογοτεχνικά αποσπάσματα. *

1 2 3 4 5
Διαφωνώ απόλυτα. ☐ ☐ ☐ ☐ ☐ Συμφωνώ απόλυτα.

8. Συνολικά, είμαι ικανοποιημένος/-η με το εργαλείο. Η εμπειρία κρίνω πως ήταν παιδαγωγικά αξιολογή. *

1 2 3 4 5
Διαφωνώ απόλυτα. ☐ ☐ ☐ ☐ ☐ Συμφωνώ απόλυτα.

Επόμενο

Εκκαθάριση φόρμας

Μέρος 2ο

Μαθησιακή εμπλοκή του χρήστη.

9. Ένιωσα ότι ο χρόνος πέρασε γρήγορα κατά τη χρήση του ψηφιακού βοηθού. *

1 2 3 4 5

Διαφωνώ απόλυτα. ☐ ☐ ☐ ☐ ☐ Συμφωνώ απόλυτα.

10. Ένιωσα ότι συμμετείχα ενεργά — δεν περίμενα απλώς να μου δώσει απαντήσεις το σύστημα. *

1 2 3 4 5

Διαφωνώ απόλυτα. ☐ ☐ ☐ ☐ ☐ Συμφωνώ απόλυτα.

11. Αισθάνθηκα ασφάλεια να κάνω ερωτήσεις που δεν θα έκανα στην τάξη. *

1 2 3 4 5

Διαφωνώ απόλυτα. ☐ ☐ ☐ ☐ ☐ Συμφωνώ απόλυτα.

12. Το σύστημα με βοήθησε να συνδέσω το κείμενο με κάτι που ήδη γνώριζα από το μάθημα. *

1 2 3 4 5

Διαφωνώ απόλυτα. ☐ ☐ ☐ ☐ ☐ Συμφωνώ απόλυτα.

?

Πίσω Επόμενο

Εκκαθάριση φόρμας

Μέρος 3ο

Ερωτήσεις ανοιχτού τύπου

13. Τι σου άρεσε περισσότερο στον τρόπο που το σύστημα σε βοήθησε να κατανοήσεις το κείμενο; *

Η απάντησή σας

14. Υπήρξε κάτι που ο ψηφιακός βοηθός έκανε λάθος ή που δεν σε βοήθησε; Περιγράψέ το. *

Η απάντησή σας

15. Πού διέφερε η βοήθεια του συστήματος από αυτή ενός καθηγητή; *

Η απάντησή σας

16. Αν μπορούσες να αλλάξεις κάτι στο εργαλείο, τι θα άλλαζες; *

Η απάντησή σας

Πίσω

Υποβολή

Εκκαθάριση φόρμας

Μην υποβάλετε ποτέ κωδικούς πρόσβασης μέσω των Φορμών Google.

Αυτό το περιεχόμενο δεν έχει δημιουργηθεί και δεν έχει εγκριθεί από την Google. - [Κάταχες φόρμες επικοινωνίας](#) - [Όροι παροχής υπηρεσιών](#) - [Πολιτική απορρήτου](#)

Αυτή η φόρμα φαίνεται ύποπτη; [Αναφορά](#)

Google Φόρμες

?

Παράρτημα III

Εντυπο αξιολόγησης καθηγητή

Συμπληρώνεται μετά τη χρήση του AI εργαλείου.
Εκτιμώμενος χρόνος: ~15 λεπτά.

fgerodim@gmail.com [Εναλλαγή λογαριασμού](#)

Δεν κοινοποιήθηκε

* Υποδεικνύει απαιτούμενη ερώτηση

1. Το σύστημα θέτει ερωτήσεις αντί να δίνει άμεσες απαντήσεις. *

1

2

3

4

5

Διαφωνώ απόλυτα.

☐

☐

☐

☐

☐

Συμφωνώ απόλυτα.

2. Σε περίπτωση λάθους, δίνεται υπόδειξη αντί της σωστής απάντησης. *

1

2

3

4

5

Διαφωνώ απόλυτα.

☐

☐

☐

☐

☐

Συμφωνώ απόλυτα.

3. Το επίπεδο καθοδήγησης είναι ανάλογο της δυσκολίας. *

1

2

3

4

5

Διαφωνώ απόλυτα.

☐

☐

☐

☐

☐

Συμφωνώ απόλυτα.

4. Παρατηρείται σταδιακή μείωση υποστήριξης (fading) κατά τη συνεδρία. *

1

2

3

4

5

Διαφωνώ απόλυτα.

☐

☐

☐

☐

☐

Συμφωνώ απόλυτα.

5. Το λεξιλόγιο και η γραμματική ταιριάζουν με το CEFR επίπεδο. *

	1	2	3	4	5	
Διαφωνώ απόλυτα.	☐	☐	☐	☐	☐	Συμφωνώ απόλυτα.

6. Οι οδηγίες των ασκήσεων είναι πλήρεις και κατανοητές για τον μαθητή. *

	1	2	3	4	5	
Διαφωνώ απόλυτα.	☐	☐	☐	☐	☐	Συμφωνώ απόλυτα.

7. Το σύστημα αναπροσαρμόζει τη δυσκολία βάσει των απαντήσεων του μαθητή. *

	1	2	3	4	5	
Διαφωνώ απόλυτα.	☐	☐	☐	☐	☐	Συμφωνώ απόλυτα.

8. Η ανατροφοδότηση αναφέρεται στη συγκεκριμένη απάντηση του μαθητή. *

	1	2	3	4	5	
Διαφωνώ απόλυτα.	☐	☐	☐	☐	☐	Συμφωνώ απόλυτα.

9. Στις απαντήσεις του συστήματος χρησιμοποιείται θετική ενίσχυση. *

	1	2	3	4	5	
Διαφωνώ απόλυτα.	☐	☐	☐	☐	☐	Συμφωνώ απόλυτα.

10. Οι πληροφορίες που σχετίζονται με το κείμενο/συγγραφέα είναι ακριβείς. *

1 2 3 4 5
Διαφωνώ απόλυτα. ☐ ☐ ☐ ☐ ☐ Συμφωνώ απόλυτα.

11. Οι ασκήσεις και τα παραδείγματα προέρχονται από το corpus. *

1 2 3 4 5
Διαφωνώ απόλυτα. ☐ ☐ ☐ ☐ ☐ Συμφωνώ απόλυτα.

Επόμενο

Εκκαθάριση φόρμας

Μέρος 2ο

Συνολική εκτίμηση συνεδρίας

12. Πως κρίνω το σύστημα ως προς την παιδαγωγική αποτελεσματικότητα. *

- ☐ Πολύ κακή
- ☐ Κακή
- ☐ Μέτρια
- ☐ Καλή
- ☐ Άριστη

13. Πώς κρίνω το σύστημα ως προς την ακρίβεια σε σχέση με το CEFR επίπεδο. *

- ☐ Πολύ κακή
- ☐ Κακή
- ☐ Μέτρια
- ☐ Καλή
- ☐ Άριστη

14. Πώς κρίνω το σύστημα ως προς την κινητροδότηση και την εμπλοκή του μαθητή. *

- ☐ Πολύ κακή
- ☐ Κακή
- ☐ Μέτρια
- ☐ Καλή
- ☐ Άριστη

Πίσω

Επόμενο

Εκκαθάριση φόρμας

Μέρος 3ο

Συνέντευξη

1. Πώς αξιολογείτε συνολικά την παιδαγωγική λογική του συστήματος; *

-Ο Σωκρατικός διάλογος ήταν γνήσιος ή μηχανιστικός;

-Σε τι διαφέρει από ένα chatbot γενικής χρήσης;

Η απάντησή σας

2. Η γλωσσική δυσκολία ήταν ευθυγραμμισμένη με το επίπεδο του μαθητή; *

- Εντοπίσατε λανθασμένες πληροφορίες ή παρερμηνείες του κειμένου;

-Ήταν η λογοτεχνική ορολογία κατάλληλη για το επίπεδο;

- Πώς χειρίστηκε τη διαφορά μεταξύ επιπέδων;

Η απάντησή σας

3. Οι μαθητές εμπλέκονταν ουσιαστικά ή απλώς συμπλήρωναν απαντήσεις; *

- Η ανατροφοδότηση ήταν εξατομικευμένη και χρήσιμη ή γενική;

- Πώς συγκρίνεται με τη δική σας ανατροφοδότηση ως καθηγητής;

Η απάντησή σας

14. Πως κρίνω το σύστημα ως προς την κινητροδότηση και την εμπλοκή του μαθητή. *

- ☐ Πολύ κακή
- ☐ Κακή
- ☐ Μέτρια
- ☐ Καλή
- ☐ Άριστη

[Πίσω](#)

[Επόμενο](#)

[Εκκαθάριση φόρμας](#)

Μην υποβάλετε ποτέ κωδικούς πρόσβασης μέσω των Φορμών Google.

Αυτό το περιεχόμενο δεν έχει δημιουργηθεί και δεν έχει εγκριθεί από την Google. - [Κάτοχος φόρμας](#)
[επικοινωνίας](#) - [Όροι παροχής υπηρεσιών](#) - [Πολιτική απορρήτου](#)

Αυτή η φόρμα φαίνεται ύποπτη; [Αναφορά](#)

Google Φόρμες

Παράρτημα IV

Chat Assistant

Target CEFR Level:

B1

Upload Text Corpus (.txt only):

Επιλογή αρχείου

Δεν επιλέχθηκε κανένα αρχείο.

This text will be sent to the AI for exercise generation.

Select an AI Assistant:

Gemini

Chat reset. Using the GEMINI model.

Reset Chat History

Υπεύθυνη Δήλωση Συγγραφέα:

Δηλώνω ρητά ότι, σύμφωνα με το άρθρο 8 του Ν.1599/1986, η παρούσα εργασία αποτελεί αποκλειστικά προϊόν προσωπικής μου εργασίας, δεν προσβάλλει κάθε μορφής δικαιώματα διανοητικής ιδιοκτησίας, προσωπικότητας και προσωπικών δεδομένων τρίτων, δεν περιέχει έργα/εισφορές τρίτων για τα οποία απαιτείται άδεια των δημιουργών/δικαιούχων και δεν είναι προϊόν μερικής ή ολικής αντιγραφής, οι πηγές δε που χρησιμοποιήθηκαν περιορίζονται στις βιβλιογραφικές αναφορές και μόνον και πληρούν τους κανόνες της επιστημονικής παράθεσης.