



ΚΟΙΝΟΤΙΚΟ ΠΛΑΙΣΙΟ ΣΤΗΡΙΞΗΣ – ΕΠΕΑΕΚ

Άξονας Προτερ.:	Πρώθηση & Βελτίωση της εκπαίδευσης και της αρχικής επαγγελματικής κατάρτισης στο πλαίσιο της διά βίου μάθησης
Μέτρο:	Αναβάθμιση της ποιότητας της παρεχόμενης εκπαίδευσης
Κωδικός Έργου:	111004
Τίτλος Έργου:	ΣΥΣΤΗΜΑ ΑΞΙΟΛΟΓΗΣΗΣ ΚΑΙ ΠΙΣΤΟΠΟΙΗΣΗΣ ΓΛΩΣΣΟΜΑΘΕΙΑΣ ΣΤΟ ΑΠΘ
Ανάδοχος Έργου:	Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Φορέας Επίβλεψης:	Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
Κωδικός Υποέργου:	02
Τίτλος Υποέργου:	<u>Ετοιμασία Οργάνων Μέτρησης για την Ιταλική Γλώσσα</u>
Υπεύθυνος Υποέργου:	Αντώνης Τσοπάνογλου (Καθηγητής της Διδακτικής των Γλωσσών, Τμήμα Ιταλικής Γλώσσας και Φιλολογίας)
Πακέτο εργασίας:	07 (Έλεγχος Ποιότητας Χρησιμοποιηθέντων Οργάνων Μέτρησης: Item Analysis)
Παραδοτέο:	12 (Αναλυτική έκθεση δεδομένων και αποτελεσμάτων)

Θεσσαλονίκη, 30 Ιουλίου 2008

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

1. ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ	3
1.1. Εισαγωγή.....	3
1.2. Έλεγχος της αξιολόγησης που έγινε βάσει κριτηρίου	3
1.3. Έλεγχος της αξιολόγησης που έγινε βάσει νόρμας.....	6
2. ΔΕΔΟΜΕΝΑ ΑΠΟ ΤΑ ΘΕΜΑΤΑ ΙΤΑΛΙΚΗΣ	6
2.1. Πρώτη ενότητα	7
2.2. Τρίτη ενότητα.....	17
3. ΔΕΙΓΜΑ ΔΕΔΟΜΕΝΩΝ ΑΠΟ ΘΕΜΑΤΑ ΑΛΛΩΝ ΓΛΩΣΣΩΝ	19
3.1. Αγγλική γλώσσα	19
3.2. Γαλλική γλώσσα	21
3.3. Γερμανική γλώσσα	23
4. ΣΥΓΚΡΙΤΙΚΗ ΑΝΑΛΥΣΗ	25
4.1. Πρώτη ενότητα	25
4.2. Τρίτη ενότητα	29
5. ΣΥΜΠΕΡΑΣΜΑΤΑ	30
ΠΑΡΑΡΤΗΜΑ: Στικτά διαγράμματα	32
Α. Μάιος 2007	32
Β. Νοέμβριος 2007	35
Γ. Μάιος 2008	38

1. ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ

1.1. Εισαγωγή

Τα χρησιμοποιηθέντα όργανα του Κρατικού Πιστοποιητικού Γλωσσομάθειας είναι αναπόσπαστο κομμάτι μιας διαδικασίας που εντάσσεται σε αυτό που λέμε «εκπαιδευτική αξιολόγηση», και που εδώ αντιμετωπίζεται ως επιστημονική έρευνα. Η επιστημονική έρευνα απαιτεί πάντα κάποιον έλεγχο της εγκυρότητάς της, η οποία εγκυρότητα περιλαμβάνει/προϋποθέτει και την αξιοπιστία του οργάνου που χρησιμοποιείται. Αυτή η δήλωση-θέση κάνει περιττή την επανάληψη από εδώ και πέρα της αξιοπιστίας ως επιθυμίας ή επιδίωξης (βλέπε κεφάλαιο Νο 2 Παραδοτέου Νο 6).

Κατά τον έλεγχο αυτού του είδους έρευνας ένα πρώτο βασικό ερώτημα που πρέπει να απαντήσει ο ερευνητής είναι αν η αξιολόγηση που τον απασχολεί έγινε με βάση κάποιο κριτήριο ή με βάση τη νόρμα. Το ερώτημα είναι σημαντικό επειδή, αν η αξιολόγηση έγινε με βάση το κριτήριο, δηλαδή τους στόχους και το περιεχόμενο ενός συγκεκριμένου προγράμματος που χρησιμοποιήθηκε από μικρό αριθμό μαθητών, τότε η κατανομή συχνοτήτων της βαθμολογίας που εξασφάλισαν οι μαθητές κατά την αξιολόγηση είναι πολύ πιθανό να μην είναι κανονική, ισόρροπη. Σε αυτή την περίπτωση ο δείκτης διάκρισης που παρουσιάζεται παρακάτω και που είναι από τους πιο βασικούς για τον έλεγχο ενός τεστ δεν έχει νόημα. Ο μόνος έλεγχος που έχει νόημα σε τέτοια περίπτωση είναι ο έλεγχος εγκυρότητας περιεχομένου (βλέπε και Παραδοτέο Νο 6, σελ. 13-4) και η σύγκριση της επίδοσης του κάθε μαθητή, στο ίδιο τεστ, στην αρχή της μαθησιακής προσπάθειας και στο τέλος της.

1.2. Έλεγχος της αξιολόγησης που έγινε βάσει κριτηρίου

Η εγκυρότητα περιεχομένου είναι, προφανώς σκόπιμο να ελέγχεται, είτε η αξιολόγηση είναι «criterion referenced», είτε «norm referenced». Επειδή, όμως, συμβαίνει κάποτε η αξιολόγηση με βάση τη νόρμα να γίνεται χωρίς σαφώς διατυπωμένο πρόγραμμα, ο έλεγχος εγκυρότητας περιεχομένου δεν είναι δυνατός, πράγμα που δεν μπορεί ποτέ να συμβαίνει αν η αξιολόγηση διεξάγεται με βάση το κριτήριο.

Υπό μορφή παραδείγματος θα αναφέρω την περίπτωση της αξιολόγησης, στο πλαίσιο των «πανελληνίων» αρχικά και «γενικών» στη συνέχεια εξετάσεων, των υποψηφίων

φοιτητών των ξενόγλωσσων τμημάτων των ΑΕΙ. Οι συγκεκριμένες εξετάσεις για την ιταλική γλώσσα (με 600 περίπου εξεταζόμενους κατά μέσο όρο κάθε χρόνο, τα τελευταία 20 χρόνια για τα οποία διαθέτω στοιχεία) διεξάγονται χωρίς πρόγραμμα, που να λείπει ποιες επικοινωνιακές δεξιότητες αυτοί πρέπει να έχουν ή, έστω, ποια μορφολογικά φαινόμενα να μπορούν να αντιμετωπίσουν. Χωρίς τέτοιο πρόγραμμα είναι αδύνατο να αποφανθεί κανείς για το αν διαθέτει, και σε ποιο βαθμό, εγκυρότητα περιεχομένου η αξιολόγηση την οποία οργανώνει το ΥΠΕΠΘ. Θα μπορούσε κάποιο σημαντικό κεφάλαιο της ιταλικής γραμματικής (για παράδειγμα) να μην εμφανιστεί ποτέ σε κανένα από τα τεστ, ενώ κάποιο άλλο, ήσσονος ίσως σημασία, να επαναλαμβάνεται κάθε χρόνο, από συνήθεια ή για οποιονδήποτε άλλο λόγο.

Αντίστοιχη κατάσταση ίσχυσε επί χρόνια και για τις εξετάσεις για τη γερμανική γλώσσα, πριν αυτή αρχίσει να διδάσκεται στη Μέση Εκπαίδευση. Για την αγγλική και γαλλική, αντίθετα, υπάρχει το αναλυτικό πρόγραμμα της Μέσης που κάνει δυνατό τον έλεγχο της εγκυρότητας περιεχομένου. Άλλωστε η κοινωνική αναστάτωση που προκαλείται όταν σε εξετάσεις πανελλαδικού χαρακτήρα, σε οποιοδήποτε μάθημα, «πέφτουν» θέματα «εκτός ύλης» είναι το αποτέλεσμα μιας μορφής υποτυπώδους ελέγχου εγκυρότητας περιεχομένου, αφού εντοπίζονται περιεχόμενα του τεστ (δηλαδή του δείγματος) που δεν υπάρχουν στο αναλυτικό πρόγραμμα (δηλαδή στον πληθυσμό).

Όταν ένας δάσκαλος αξιολογεί με βάση το κριτήριο, όπως ήδη δηλώθηκε, θα έχει προφανώς πρόγραμμα και θα μπορεί να ελέγχει την εγκυρότητα περιεχομένου. Ένα δεύτερο πράγμα που μπορεί να ελέγξει είναι, με την τεχνική των ημιμορίων, η αξιοπιστία του τεστ. Αυτό, όμως, έχει ως προϋπόθεση πως το τεστ αποτελείται από ερωτήματα ίδιου ειδικού βάρους (πράγμα που στο ΚΠΓ ισχύει μόνο στις ενότητες 1 και 3), τουλάχιστον όπως αυτό εκφράζεται με την κλίμακα βαθμολόγησης. Αν η σωστή απάντηση σε κάποιο ερώτημα εξασφαλίζει μία μονάδα (δηλαδή βαθμολογείται με δίβαθμη κλίμακα: 0 έως 1) και η επόμενη εξασφαλίζει πέντε μονάδες (εξάβαθμη κλίμακα: 0 έως 5), η τεχνική αυτή δεν είναι η κατάλληλη.

Ο ιδανικός, ίσως, τρόπος ελέγχου του οργάνου στην περίπτωση αξιολόγησης με βάση το κριτήριο είναι το ίδιο ακριβώς τεστ να δοθεί στην αρχή και στο τέλος της μαθησιακής προσπάθειας στους ίδιους ακριβώς μαθητές. Η αυτονόητη παραδοχή είναι πως οι μαθητές αρχίζουν μια σειρά μαθημάτων επειδή θέλουν/χρειάζεται να μάθουν «κάτι» και τελειώνουν τον κύκλο μαθημάτων έχοντας όλοι, σε διαφορετικό ίσως βαθμό, μάθει αυτό το «κάτι». Αν κατασκευάσει ο δάσκαλος-ερευνητής ένα όργανο που να μετράει την κατοχή από τα υποκείμενα αυτού του «κάτι» και να μη μετράει οτιδήποτε άλλο, τότε η αρχική και η τελική μέτρηση πρέπει να δώσουν σημαντικά διαφορετικές τιμές. Αν αυτό δε συμβεί, τότε πρέπει να ελέγξουμε μήπως: α) η διδασκαλία-μάθηση απέτυχε, β) τα διδακτικά υλικά ήταν ακατάλληλα, γ) η αξιολογική έρευνα δεν είχε εγκυρότητα, δ) οι μαθητές αντιστέκονται στην εκμάθηση της συγκεκριμένης ύλης ή στην απόκτηση της

συγκεκριμένης δεξιότητας (που αποτελούσαν το στόχο του μαθήματος) για πολιτισμικούς ή άλλους λόγους, κτλ.

Ο έλεγχος της διαφοράς μεταξύ ελέγχου και μετελέγχου γίνεται υπολογίζοντας τον απλούστατο στατιστικό δείκτη που είναι γνωστός ως δείκτης διαφοράς (difference index). Για τον υπολογισμό του χρειάζεται πρώτα να βρεθεί ο δείκτης ευκολίας (παρουσιάζεται παρακάτω) που είχε το ερώτημα κατά την αρχική χρησιμοποίηση του τεστ με τα υποκείμενα και ο δείκτης ευκολίας του ίδιου ερωτήματος κατά τον μετέλεγχο, με τα ίδια υποκείμενα. Ο δείκτης διαφοράς προκύπτει από την αφαίρεση του δείκτη ευκολίας του κάθε ερωτήματος κατά τον έλεγχο και κατά τον μετέλεγχο.

Σε περίπτωση που, για ένα ερώτημα του τεστ, ο δείκτης διαφοράς δεν είναι αρκετά υψηλός, και για να αποκλειστεί καθεμία από τις υποθέσεις που καταγράφονται στην προτελευταία παράγραφο και να βεβαιωθούμε ότι το τεστ μετρούσε πράγματι αυτό που θέλαμε να μετράει, είναι σκόπιμο να ελεγχθεί η συνυπάρχουσα εγκυρότητά (concurrent validity) του. Αυτό είναι δυνατό αν χρησιμοποιηθεί με υποκείμενα που αποδεδειγμένα διαθέτουν το κάτι που το τεστ υποτίθεται ότι μετράει.

Αν είναι πρακτικά αδύνατο να δώσει κανείς το ίδιο τεστ δύο φορές στα ίδια άτομα, πριν και μετά τη μεσολάβηση μάθησης, τότε είναι δυνατό να υπολογίσει τη διαφορά μεταξύ βαθμού ευκολίας του κάθε ερωτήματος χωριστά στην ομάδα των υποκειμένων που πήραν «προβιβασμό» βαθμό στο τεστ και στην ομάδα που θεωρήθηκε πως δεν ικανοποιούσε το κριτήριο, δηλαδή δεν πέτυχε σε ικανοποιητικό βαθμό τους επιδιωκόμενους μαθησιακούς στόχους. Πρόκειται ουσιαστικά για ένα «υποκατάστατο» του δείκτη διαφοράς, και βγαίνει πάλι με μίαν απλή αφαίρεση. Ο δείκτης αυτός, που είναι γνωστός στην αγγλική ως B-index είναι ταυτόχρονα και μια μορφή του δείκτη διάκρισης που θα εξεταστεί λεπτομερειακά παρακάτω.

Συμπερασματικά μπορεί να πει κανείς ότι, ως προς την αξιολόγηση βάσει κριτηρίου, ο έλεγχος εγκυρότητας είναι εφικτός χωρίς να απαιτούνται ειδικές τεχνικές γνώσεις. Αρκεί η λογική σκέψη και η συστηματική διερεύνηση υποθέσεων. Ο Popham, που έχει ειδικά ασχοληθεί με την αξιολόγηση βάσει κριτηρίου, θεωρεί πως ο έλεγχος διευκολύνεται πολύ αν για κάθε ερώτημα ή κάθε ομάδα ερωτημάτων (δοκιμασία) ο συντάκτης ενός τεστ έχει καταγράψει: α) μια σύντομη περιγραφή του μαθησιακού στόχου με τον οποίο αυτό συνδέεται, β) ένα δείγμα ερωτήματος, γ) τα χαρακτηριστικά του ερεθίσματος που θεωρεί πως πρέπει να δοθεί στον αξιολογούμενο, για παράδειγμα το κείμενο που θα αποτελεί το εισιόν (input), δ) τα χαρακτηριστικά της αναμενόμενης από τον αξιολογούμενο αντίδρασης, δηλαδή το εξιόν (output) και ε) ενδεχόμενα βοηθήματα (Popham, 1981). Όταν αυτός που ελέγχει την εγκυρότητα μιας εξέτασης, είτε είναι το ίδιο άτομο που την οργάνωσε είτε, ακόμη περισσότερο, είναι άλλος ερευνητής, διαθέτει τις παραπάνω πληροφορίες διευκολύνεται πολύ στο έργο του.

1.3. Έλεγχος της αξιολόγησης που έγινε βάσει νόρμας

Οι διαδικασίες ελέγχου της εγκυρότητας σε περίπτωση αξιολόγησης μαθητών σε ευρεία κλίμακα και με βάση τη νόρμα εξηγούνται διεξοδικά στο κεφάλαιο Νο 4 του Παραδοτέου Νο 6, στο οποίο παραπέμπεται ο αναγνώστης

2. ΔΕΔΟΜΕΝΑ ΑΠΟ ΤΑ ΘΕΜΑΤΑ ΙΤΑΛΙΚΗΣ

Ακολουθώντας τις βασικές αρχές που εκτέθηκαν παραπάνω καθώς και, κυρίως, στο Παραδοτέο Νο 6 έγινε, στο πλαίσιο του παρόντος Πακέτου Εργασίας, ο βασικός έλεγχος της ποιότητας των ολοκληρωμένων οργάνων μέτρησης της ιταλομάθειας που κατασκευάστηκαν.

Χρησιμοποιήθηκαν και άλλες τεχνικές στις οποίες δεν κρίθηκε απαραίτητο να γίνει λεπτομερής αναφορά σε αυτή την έκθεση.

Τα δεδομένα αυτού του κεφαλαίου αφορούν μόνο την ιταλική γλώσσα, αλλά σε επόμενο κεφάλαιο θα παρουσιαστούν μερικά δεδομένα από θέματα εξέτασης που ετοιμάστηκαν από αντίστοιχα Υποέργα για τις γλώσσες αγγλική, γαλλική και γερμανική, έτσι ώστε, στο τέταρτο κεφάλαιο της αναφοράς/έκθεσης να είναι δυνατή η σύγκριση των δεδομένων και η εξαγωγή συμπερασμάτων ως προς την ποιότητα των οργάνων ιταλικής. Αυτό κρίθηκε χρήσιμο επειδή οποιοσδήποτε έλεγχος ποιότητας είναι μια μορφή αξιολόγησης, και η βασική νοητική διεργασία για την οποιαδήποτε αξιολόγηση είναι η **σύγκριση**.

2.1. Πρώτη ενότητα

Ο έλεγχος των ολοκληρωμένων οργάνων μέτρησης που είχαμε την υποχρέωση να ετοιμάσουμε έγινε, κατά την περίοδο Μαΐου 2007 πάνω σε δείγμα 2.149 ατόμων. Το πλήθος των 3.006 ατόμων που εμφανίζονται στον πίνακα της επόμενης σελίδας είναι οι υποψήφιοι. Οι αριθμοί που θα δίνονται στο εξής θα είναι οι αριθμοί των υποψηφίων και όχι αυτών που πραγματικά εξετάστηκαν. Κι αυτό επειδή, ενώ μου είναι γνωστή η γεωγραφική κατανομή των υποψηφίων δεν είναι γνωστή η γεωγραφική κατανομή των εξετασθέντων ανά την επικράτεια. Όπως προκύπτει από τον παρακάτω πίνακα, η διαφορά μεταξύ υποψηφίων και προσελθόντων στις δύο περιόδους του 2007 είναι της τάξεως του 22,65%. Με άλλα λόγια, 941 άτομα (από τα 4.154) δήλωσαν ότι θέλουν να κάνουν την εξέταση στην ιταλική, πλήρωσαν το σχετικό παράβολο και δεν ήρθαν καθόλου στο εξεταστικό κέντρο. Πρόκειται για σημαντικό ποσοστό, που όμως δεν δημιουργεί πρόβλημα, αφού στη βιβλιογραφία αναφέρεται πως η στατιστική επεξεργασία που κάναμε δε θεωρείται χρήσιμη αν τα υποκείμενα είναι λιγότερα από 100, πράγμα που δε συνέβη με κανένα επίπεδο της ιταλικής και σε καμία εξεταστική περίοδο. Στον πίνακα της επόμενης σελίδας φαίνεται η γεωγραφική κατανομή των ατόμων.

Εξεταστική περίοδος	Επίπεδο	Υποψήφιοι	Προσελθόντες
ΜΑΪΟΣ 2007	B1	608	533
	B2	2.015	1.804
	Γ1	384	345
	ΣΥΝΟΛΟ	3.007	2.149
ΝΟΕΜΒΡΙΟΣ 2007	B1	251	227
	B2	706	651
	Γ1	190	186
	ΣΥΝΟΛΟ	1.147	1.064

Κατά την εξεταστική περίοδο Νοεμβρίου 2007 το δείγμα ατόμων ήταν 1.147 άτομα (βλέπε επόμενη σελίδα), των οποίων η γεωγραφική κατανομή δεν μας είναι ακριβώς γνωστή, γνωρίζουμε, ωστόσο, ότι αναλογικά είναι ίδια με αυτήν της περιόδου Μαΐου.

Τέλος, κατά την εξεταστική περίοδο Μαΐου 2008 διαθέταμε δείγμα 2.787 ατόμων (βλέπε επόμενη σελίδα), των οποίων τη γεωγραφική κατανομή γνωρίζουμε -και μάλιστα ανά εξεταστικό κέντρο, δηλαδή με πολύ μεγάλη ακρίβεια- αλλά για λόγους οικονομίας δε θεωρούμε σκόπιμο να δώσουμε εδώ. Έχουμε, πάντως, και πάλι την ίδια περίπου, αναλογικά, κατανομή στην επικράτεια.

ΚΑΤΑΝΟΜΗ ΕΞΕΤΑΣΘΕΝΤΩΝ ΠΕΡΙΟΔΟΥ ΜΑΪΟΥ 2007													
Περιφέρεια	Γλώσσα/Επίπεδο												ΣΥΝ.
	En B1	En B2	En Γ1	Fr B1	Fr B2	Fr Γ1	DE B1	De B2	De Γ1	It B1	It B2	It Γ1	
ΑΝΑΤΟΛΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ ΚΑΙ ΘΡΑΚΗΣ	73	443	183	16	36	12	48	150	38	25	101	14	1139
ΑΤΤΙΚΗΣ	234	1552	928	46	150	63	71	272	68	115	574	149	4222
ΒΟΡΕΙΟΥ ΑΙΓΑΙΟΥ	84	279	138	16	39	14	20	40	5	13	14	8	670
ΔΥΤΙΚΗΣ ΕΛΛΑΔΟΣ	36	600	157	13	43	10	13	85	6	81	161	27	1232
ΔΥΤΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ	25	546	179	5	36	24	20	63	22	29	70	15	1034
ΗΠΕΙΡΟΥ	36	284	153	23	60	10	51	151	19	53	88	21	949
ΘΕΣΣΑΛΙΑΣ	94	771	398	10	75	11	24	70	28	23	155	29	1688
ΙΟΝΙΩΝ ΝΗΣΩΝ	20	198	53	5	22	2	3	41		27	50	6	427
ΚΕΝΤΡΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ	210	1807	749	50	144	36	95	242	65	103	515	58	4074
ΚΡΗΤΗΣ	42	430	165	11	59	9	76	146	27	58	124	25	1172
ΝΟΤΙΟΥ ΑΙΓΑΙΟΥ	74	273	102	6	3	2	54	54	18	7	21	6	620
ΠΕΛΟΠΟΝΝΗΣΟΥ	49	290	164	20	103	8	16	44	7	57	92	16	866
ΣΤΕΡΕΑΣ ΕΛΛΑΔΑΣ	34	574	174	17	91	11	5	35	9	17	50	9	1026
ΣΥΝΟΛΑ	1011	8047	3543	238	861	212	496	1393	312	608	2015	383	19119
ΣΥΝΟΛΑ ΑΝΑ ΓΛΩΣΣΑ	12601			1311			2201			3006			

ΚΑΤΑΝΟΜΗ ΕΞΕΤΑΣΘΕΝΤΩΝ ΠΕΡΙΟΔΟΥ ΝΟΕΜΒΡΙΟΥ 2007				
	ΑΓΓΛΙΚΗ	ΓΑΛΛΙΚΗ	ΓΕΡΜΑΝΙΚΗ	ΙΤΑΛΙΚΗ
ΕΠΙΠΕΔΟ Β1	377	220	520	251
ΕΠΙΠΕΔΟ Β2	3.452	584	724	706
ΕΠΙΠΕΔΟ Γ1	1.173	123	201	190
ΣΥΝΟΛΟ	5.002	927	1.445	1.147

ΚΑΤΑΝΟΜΗ ΕΞΕΤΑΣΘΕΝΤΩΝ ΠΕΡΙΟΔΟΥ ΜΑΪΟΥ 2008																	
	Γλώσσα/Επίπεδο														ΣΥΝ.		
	En A	En B1	En B2	En Γ1	Fr A	Fr B1	Fr B2	Fr Γ1	De A	De B1	De B2	De Γ1	It A	It B1		It B2	It Γ1
ΣΥΝ.	426	957	7.239	2.814	296	529	1.047	301	412	935	1.141	412	240	512	1.691	344	19.296
	11.436				2.173				2.900				2.787				

Όπως προκύπτει από τους παραπάνω πίνακες, ο αριθμός των ατόμων που είχαμε στη διάθεσή μας για να ελέγξουμε a posteriori τα όργανα μέτρησης του ΚΠΓ αποτελούσαν κάθε φορά ένα δείγμα καθόλα ικανοποιητικό: και ποσοτικά και ποιοτικά, δηλαδή ως προς την αντιπροσωπευτικότητα, αφού υπήρξε κατανομή στο χώρο ανάλογη με τον πληθυσμό. Αξίζει, ωστόσο, να παρατηρήσουμε δύο πράγματα: α) η περίοδος Νοεμβρίου

είχε πολύ μικρότερο αριθμό υποκειμένων σε σχέση με τις δύο περιόδους Μαΐου και β) οι υποψήφιοι για το επίπεδο Γ1 ήταν κάθε φορά πολύ λιγότεροι από τους υποψήφιους των άλλων επιπέδων. Κανένα από τα δύο παραπάνω δεδομένα δε θέτει εν αμφιβόλω την έρευνα που έγινε και που παρουσιάζεται παρακάτω, αφού οι αριθμοί αυτοί είναι επαρκείς για εξαγωγή ασφαλών συμπερασμάτων.

Στη συνέχεια παρουσιάζονται τα αποτελέσματα της ανάλυσης ερωτημάτων (item analysis) τα οποία προέκυψαν από την εφαρμογή των στατιστικών τύπων που παρουσιάστηκαν στο Παραδοτέο Νο 6.

Θα περιοριστούμε στην παρουσίαση και σχολιασμό των δεδομένων που προκύπτουν από τον έλεγχο των θεμάτων περιόδου Μαΐου 2007 για λόγους οικονομίας. Στο συγκριτικό έλεγχο των θεμάτων σε σχέση με αυτά και των άλλων γλωσσών θα εμφανιστεί το αποτέλεσμα, δηλαδή η τελική αποτίμηση της ποιότητας των οργάνων που ετοιμάσαμε και για τις άλλες δύο εξεταστικές περιόδους.

Για οικονομία χώρου, από τον πίνακα που ακολουθεί και που παρουσιάζει τους δείκτες ευκολίας και διάκρισης των ερωτημάτων επιπέδου Β1, της πρώτης ενότητας (κατανόηση γραπτού λόγου) της περιόδου Μαΐου 2007 έχουν αφαιρεθεί τα ερωτήματα 42-46, των οποίων η ποιότητα απεικονίζεται, ωστόσο στα νούμερα 32,0 και 51,5, που είναι ο μέσος όρος των δεικτών ευκολίας και διάκρισης όλων των ερωτημάτων της έκτης δοκιμασίας του τεστ. Μετά από τους τρεις πίνακες ο αναγνώστης του παρόντος Παραδοτέου βρίσκει διευκρινίσεις αναφορικά με τα χαρακτηριστικά τους.

ΜΑΙΟΣ 2007 Β1

	e.i.*100	d.i.*100	A	B	C	D	E	F	G	H	χ	
1	53	53	8	19	0	6	11	2	53	1	1	
2	89	30	1	2	0	2	1	3	2	89	0	
3	83	25	1	1	1	83	3	2	3	5	0	
4	84	33	1	0	1	5	4	84	5	1	0	
5	95	15	1	0	95	0	2	0	1	0	0	
6	74	39	3	74	0	0	4	1	15	1	1	1
7	65	51	65	2	4	1	10	2	13	2	1	76,0
8	65	51	19	2	1	3	65	4	6	0	1	37,1
9	81	31	12	81	6						0	
10	89	10	6	4	89						1	
11	66	11	66	24	9						0	
12	58	40	19	58	23						0	
13	90	21	90	4	5						0	
14	77	43	18	4	77						0	
15	53	47	53	24	23						0	2
16	79	38	14	7	79						1	76,9
17	99	3	1	99	0						0	27,1
18	61	51	61	32	7						1	
19	70	21	70	12	17						0	
20	95	6	4	95	1						0	

21	85	15	2	13	85						0		
22	67	47	26	67	7						0		
23	66	49	30	4	66						0	3	
24	28	59	40	31	28						1	68,3	
25	74	42	74	10	16						0	36,3	
26	61	31	7	61	32						0		
27	77	38	77	11	12						0		
28	85	31	11	5	85						0		
29	60	25	38	60	1						1		
30	68	27	21	11	68						0		
31	49	55	43	49	7						1	4	
32	53	51	25	21	53						1	65,4	
33	70	39	9	70	21						1	37,1	
34	88	34	1	3	2	88	2	1	2		1		
35	78	57	8	4	3	1	78	3	2		1		
36	81	51	81	7	2	1	3	4	2		0		
37	82	44	1	2	82	1	2	0	12		1		
38	76	63	3	2	6	5	4	2	76		2	5	
39	88	29	2	1	1	0	6	88	2		1	82,3	
40	83	47	3	83	3	3	2	2	3		1	46,4	
41	25	31											
47	56	74											
48	15	35										6	
49	42	65										32,0	
50	62	60										51,5	
	65,30	39,36	average										
	22,44	16,62	st.dev.										

ΜΑΙΟΣ 2007 B2

	e.i.*100	d.i.*100	A	B	C	D	E	F	G	H	x	
1	52	19	30	18	52						1	
2	58	48	58	27	15						0	
3	41	48	38	21	41						0	
4	68	57	19	68	13						0	
5	83	38	7	83	9						0	
6	27	28	54	27	18						0	
7	87	16	2	11	87						0	
8	61	32	15	61	24						0	
9	84	27	84	3	12						0	
10	93	5	6	1	93						0	
11	76	52	76	20	3						0	
12	77	34	13	9	77						0	
13	42	53	42	32	26						1	
14	93	18	5	93	2						0	
15	18	6	53	18	28						1	1
16	82	23	11	6	82						0	65,5
17	71	52	12	71	17						0	32,7
18	73	46	73	19	8						0	
19	60	35	26	14	60						0	
20	64	37	64	18	18						0	
21	59	60	4	37	59						0	
22	49	36	35	49	16						0	

23	44	34	35	44	21						0	
24	29	40	29	37	34						0	
25	63	52	31	6	63						0	
26	66	39	20	66	13						0	
27	57	37	17	57	26						0	
28	33	51	47	20	33						0	2
29	60	42	60	34	6						0	55,3
30	62	54	62	8	31						0	43,3
31	95	12	1	0	0	1	0	95	1	0	0	
32	96	13	0	0	0	1	96	0	0	1	0	
33	90	28	90	1	1	4	0	1	0	3	0	
34	87	31	3	1	6	1	0	0	1	87	1	
35	87	36	1	1	87	3	1	0	1	5	0	
36	94	15	1	1	1	2	0	0	94	1	1	3
37	93	23	0	93	3	2	0	0	1	1	0	90,9
38	85	44	3	2	3	85	2	2	2	1	0	25,3
39	82	22	82	13	5						1	
40	69	40	69	21	10						1	
41	42	0	42	46	12						1	
42	68	14	17	15	68						1	
43	61	44	9	29	61						1	
44	44	20	44	36	19						1	
45	37	23	34	37	28						1	
46	68	53	68	17	14						1	
47	69	45	24	69	7						1	
48	31	2	29	31	39						1	4
49	47	29	7	47	46						1	54,5
50	36	39	42	21	36						1	27,6
51	26	22	22	52	26						0	
52	41	25	31	41	28						1	
53	28	4	55	16	28						1	
54	13	3	13	45	41						0	
55	22	18	46	31	22						0	
56	35	27	37	35	28						0	
57	51	20	29	51	19						0	
58	42	26	42	5	52						1	5
59	44	36	44	21	35						0	34,4
60	42	37	38	42	20						0	21,8
61	6	13										
73	40	59										6
74	4	13										21,5
75	26	48										38,9
	51,73	32,72	average									
	25,95	16,53	st.dev.									

ΜΑΙΟΣ 2007 Γ1

	e.i.*100	d.i.*100	A	B	C	D	E	F	G	H	x	
1	62	48	1	1	2	10	62	15	1	6	1	
2	57	57	4	3	7	6	5	12	4	57	2	
3	84	28	84	0	7	2	0	1	0	4	2	
4	88	26	0	0	1	0	7	1	88	2	1	
5	52	59	2	6	2	8	14	52	2	13	1	

6	63	55	2	1	63	15	5	5	1	7	2	1
7	54	55	5	2	13	54	3	10	1	8	3	68,5
8	88	24	1	88	3	2	2	1	0	2	1	44,0
9	62	44	21	62	15						1	
10	54	45	54	26	19						1	
11	51	41	41	7	51						1	
12	86	14	3	11	86						1	
13	38	38	37	24	38						1	
14	73	21	14	73	12						1	
15	37	64	37	41	21						1	
16	70	24	70	21	7						1	
17	39	38	24	39	37						1	
18	66	11	20	14	66						1	
19	46	16	23	46	30						1	
20	61	6	61	9	29						1	
21	64	37	16	64	19						1	2
22	85	20	3	85	11						1	58,7
23	49	28	49	20	30						1	29,8
24	78	31	1	1	78	5	0	11	0	2	1	
25	61	44	26	3	1	0	1	0	7	61	1	
26	72	35	3	4	10	5	1	72	2	1	2	
27	72	53	8	72	1	1	3	0	8	7	1	
28	87	34	2	5	0	0	87	0	3	1	1	
29	62	57	10	5	0	1	2	10	62	7	3	
30	50	56	50	8	1	1	4	2	15	15	3	
31	84	24	1	0	8	84	0	2	1	1	3	
32	74	34	74	17	8						1	
33	80	32	5	14	80						1	
34	73	33	16	10	73						1	
35	88	15	3	88	9						1	
36	43	38	20	43	36						1	3
37	75	43	75	9	15						1	72,9
38	94	16	2	94	3						1	36,3
39	48	36	29	48	18						6	
40	35	36	35	37	23						5	
41	59	56	59	23	14						3	
42	70	46	7	19	70						3	
43	60	26	60	15	18						7	
44	73	37	13	73	10						3	
45	34	32	32	34		29					6	
46	44	45	44	18		31					7	
47	56	41			6	56	33				5	
48	52	38			7	35	52				6	4
49	38	56			11	38	44				6	54,1
50	80	21			4	80	9				7	39,2
51	41	36	23	34	41						2	
52	86	27	86	7	5						1	
53	69	36	21	8	69						2	
54	73	39	9	73	17						2	
55	74	27	3	21	74						1	
56	60	46	19	60	20						2	
57	51	47	51	21	26						2	
58	66	19	66	24	7						3	5
59	55	44	55	17	26						2	64,0

60	65	48	9	24	65						2	36,9
61	24	24										
73	53	34										6
74	68	14										34,3
75	50	71										35,1
	57,68	36,12	average									
	20,27	14,52	st.dev.									

Οι πίνακες αυτοί έχουν 13 στήλες ο καθένας.

Στην **πρώτη στήλη** εμφανίζεται ο αύξων αριθμός του ερωτήματος. Η μαύρη γραμμή που χωρίζει κάποια από τα ερωτήματα (για παράδειγμα το όγδοο από το ένατο ερώτημα του τρίτου πίνακα (επίπεδο Γ1) φανερώνει αλλαγή στη δοκιμασία. Ο όρος δοκιμασία έχει ήδη οριστεί σε προηγούμενο Παραδοτέο.


Η **δεύτερη στήλη** καταγράφει τον δείκτη ευκολίας κάθε ερωτήματος πολλαπλασιασμένο επί 100. Δηλαδή, όπως έχει ήδη διευκρινιστεί σε προηγούμενο Παραδοτέο, οι δείκτες παίρνουν τιμή από 0 έως 1. Ο πολλαπλασιασμός με το 100 επιτρέπει την εξαφάνιση του μηδενός και του κόμματος, για οικονομία. Συνεπώς, στο εβδομηκοστό πέμπτο ερώτημα του Γ1 ο αριθμός 50 σημαίνει δείκτης ευκολίας (easiness index: e.i.) 0,50. Πράγμα που σημαίνει ότι 50% των εξετασθέντων απάντησαν σωστά στο ερώτημα.

Στην **τρίτη στήλη** εμφανίζεται ο δείκτης διάκρισης, που είναι ο σημαντικότερος (ποιοτικός δείκτης), και πάλι πολλαπλασιασμένος επί 100, για τον ίδιο λόγο.

Οι **στήλες 4-11** εμφανίζουν το ποσοστό των εξετασθέντων που επέλεξε μία από τις απαντήσεις Α έως Η που επιτρέπει το απαντητικό έντυπο αυτής της ενότητας, του οποίου την μορφή βρίσκει κανείς στην επόμενη σελίδα.

Στις στήλες αυτές ορισμένα νούμερα έχουν ανοιχτό πράσινο φόντο και ορισμένα άλλα βαθύ πράσινο. Τα πρώτα είναι τα ποσοστά επιλογής των «παραπλανητών» (distractors), δηλαδή των παραπλανητικών δυνατών απαντήσεων, και τα δεύτερα νούμερα είναι τα ποσοστά των εξετασθέντων που επέλεξαν την απάντηση που οι συντάκτες των θεμάτων θεωρούν σωστή, αναμενόμενη. Είναι αυτονόητο ότι όταν το ποσοστό της αναμενόμενης απάντησης είναι μικρότερο από το ποσοστό κάποιου παραπλανητή κάποιο σοβαρό πρόβλημα υπήρχε με το ερώτημα: ή η σωστή απάντηση ήταν υπερβολικά δύσκολη για μαθητές του συγκεκριμένου επιπέδου ή κάποιος παραπλανητής ήταν υπερβολικά «πονηρός», δηλαδή πολύ κοντά στο σωστό. Εάν, επιπλέον, ο παραπλανητής ξεγέλασε περισσότερους «καλούς» μαθητές, δηλαδή εξετασθέντες που σε όλο το τεστ έβγαλαν βαθμό που τους κατατάσσει στο 25% των καλύτερων, τότε το ερώτημα είναι κακής ποιότητας και αυτό απεικονίζεται στο δείκτη διάκρισης.

27 2-08-13 32 GRAFO A.V.E.E
*3018 4818332 # 5/ 9



**ΚΩΔΙΚΟΣ ΕΝΤΥΠΟΥ
ΓΙΑ ΤΟ ΒΑΘΜΟΛΟΓΙΚΟ ΚΕΝΤΡΟ**

Στοιχεία υποψηφίου
(γράφονται αλογράφως)

**ΘΕΣΗ ΕΤΙΚΕΤΑΣ ΜΕ ΤΟΝ
ΚΩΔΙΚΟ ΑΡΙΘΜΟ ΥΠΩΨΗΦΙΟΥ**

ΕΠΙΠΕΔΟ A1 & A2 ΕΝΤΥΠΟ ΕΝΟΤΗΤΑΣ 1 ΚΑΤΑΝΟΗΣΗ ΓΡΑΠΤΟΥ ΛΟΓΟΥ

• Στα ερωτήματα 1 έως 20 & 26 έως 45 απαντάτε μαρκίζοντας το κουτάκι έτσι —
• Στα ερωτήματα 21 έως 25 & 46 έως 50 απαντάτε γράφοντας μόνο εκεί όπου υπάρχουν τελείες.
Άλλες οδηγίες: Βλέπε πίσω σελίδα

1 A B C D E F G H

2 A B C D E F G H

3 A B C D E F G H

4 A B C D E F G H

5 A B C D E F G H

6 A B C D E F G H

7 A B C D E F G H

8 A B C D E F G H

9 A B C D E F G H

10 A B C D E F G H

11 A B C D E F G H

12 A B C D E F G H

13 A B C D E F G H

14 A B C D E F G H

15 A B C D E F G H

16 A B C D E F G H

17 A B C D E F G H

18 A B C D E F G H

19 A B C D E F G H

20 A B C D E F G H

26 A B C D E F G H

27 A B C D E F G H

28 A B C D E F G H

29 A B C D E F G H

30 A B C D E F G H

31 A B C D E F G H

32 A B C D E F G H

33 A B C D E F G H

34 A B C D E F G H

35 A B C D E F G H

36 A B C D E F G H

37 A B C D E F G H

38 A B C D E F G H

39 A B C D E F G H

40 A B C D E F G H

41 A B C D E F G H

42 A B C D E F G H

43 A B C D E F G H

44 A B C D E F G H

45 A B C D E F G H

ΠΡΟΣΟΧΗ
Συμπληρώνεται
ΜΟΝΟ
από τον βαθμολογητή

ΠΡΟΣΟΧΗ
Συμπληρώνεται
ΜΟΝΟ
από τον βαθμολογητή

21		<input type="checkbox"/> Σ <input type="checkbox"/> Λ
22		<input type="checkbox"/> Σ <input type="checkbox"/> Λ
23		<input type="checkbox"/> Σ <input type="checkbox"/> Λ
24		<input type="checkbox"/> Σ <input type="checkbox"/> Λ
25		<input type="checkbox"/> Σ <input type="checkbox"/> Λ

46		<input type="checkbox"/> Σ <input type="checkbox"/> Λ
47		<input type="checkbox"/> Σ <input type="checkbox"/> Λ
48		<input type="checkbox"/> Σ <input type="checkbox"/> Λ
49		<input type="checkbox"/> Σ <input type="checkbox"/> Λ
50		<input type="checkbox"/> Σ <input type="checkbox"/> Λ

© 37.485

Η δωδέκατη στήλη, όπου στην κεφαλή υπάρχει «x», δείχνει το ποσοστό των ατόμων που δεν απάντησαν καθόλου ή έδωσαν περισσότερες από μία απαντήσεις, πράγμα μη επιτρεπτό. Είναι προφανές ότι όταν αυτό το ποσοστό υπερβαίνει κάποιο όριο, για παράδειγμα το 3%, υπήρξε κάποιο πρόβλημα στο συγκεκριμένο ερώτημα. Συνήθως οι δοκιμα-

σίες που εμφανίζουν αυτό το πρόβλημα είναι οι δοκιμασίες «σύζευξης» ή «αντιστοίχισης» (matching).

Η δέκατη τρίτη, δηλαδή η **τελευταία, στήλη** εμφανίζει σε πορτοκαλί φόντο τον αύξοντα αριθμό της δοκιμασίας και σε μπλε φόντο τη μέση τιμή του δείκτη ευκολίας των ερωτημάτων της συγκεκριμένης δοκιμασίας (επάνω) και τη μέση τιμή του δείκτη διάκρισης της δοκιμασίας (κάτω). Υπενθυμίζεται αυτό που έχει δηλωθεί σε προηγούμενο Παραδοτέο: όσο μεγαλύτερος είναι ο δείκτης διάκρισης τόσο καλύτερη ήταν μια δοκιμασία, διότι βοήθησε στο να διακρίνουμε, να διαφοροποιήσουμε τους καλούς από τους κακούς μαθητές ή, καλύτερα, ομιλητές της ιταλικής, στην προκειμένη περίπτωση, γλώσσα. Έτσι με μια μόνο ματιά, χάρη στην τελευταία στήλη, έχουμε την πληροφορία πως η καλύτερη δοκιμασία του τεστ για το Γ1 ήταν η πρώτη, με δείκτη διάκρισης 0,44.

Στο Παραδοτέο Νο 6 είχε δοθεί ένας πίνακας με τις τιμές του δείκτη διάκρισης που φανερώνουν τότε ένα ερώτημα ή μία δοκιμασία ή ένα ολόκληρο τεστ ήταν κακό, τότε μέτριας ποιότητας και τότε καλό, ανάλογα με τον τύπο ή φόρμουλα που χρησιμοποιήθηκε. Τον αντιγράψω εδώ, με τη διευκρίνιση ότι για τον έλεγχο της ποιότητας των ολοκληρωμένων οργάνων μέτρησης χρησιμοποιήσαμε τον πρώτο τύπο εξαγωγής του δείκτη διάκρισης (γραμμή «α») του πίνακα.

Με βάση την κλείδα ερμηνείας των αποτελεσμάτων συνάγεται, για παράδειγμα, ότι όλες οι δοκιμασίες του Γ1 της περιόδου Μαΐου 2007 μπορούν να θεωρηθούν καλές, αλλά η πρώτη ήταν η καλύτερη από όλες και η δεύτερη η χειρότερη, με μέσο δείκτη διάκρισης 0,298 (29,8 στον πίνακα).

Τύπος	Ελάχιστη και μέγιστη τιμή	Ερώτημα κακό, προς αποκλεισμό	Ερώτημα μέτριο, προς βελτίωση	Ερώτημα καλό, προς επαναχρησιμοποίηση
α	-1,00 έως +1,00	-1,00 έως 0,00	+0,01 έως +0,25	+0,26 έως +1,00
β	όπως παραπ.	όπως παραπάνω	όπως παραπάνω	όπως παραπάνω
γ	-0,50 έως +0,50	-0,50 έως 0,00	+0,01 έως +0,20	+0,21 έως +0,50
δ	0,00 έως 1,00	0,00 έως 0,50	0,51 έως 0,65	0,66 έως 1,00

Οι τελευταίες δύο γραμμές της δεύτερης και της τρίτης στήλης καταγράφουν τη μέση τιμή ή αριθμητικό μέσο (average ή arithmetic mean) όλων των δεικτών ευκολίας και όλων των δεικτών διάκρισης (επάνω/προτελευταία γραμμή), καθώς και την τυπική απόκλιση (standard deviation) όλων των δεικτών ευκολίας και διάκρισης. Αυτά τα νούμερα αποτελούν το «απόσταγμα» της όλης διεργασίας και επιτρέπουν να σχηματίσει κανείς εικόνα της ποιότητας των οργάνων μέτρησης. Με χρήση αυτών των τιμών θα κάνουμε τη σύγκριση των θεμάτων στις τέσσερις γλώσσες (αγγλική, γαλλική, γερμανική και ιταλική) στο τέταρτο κεφάλαιο.

Ενώ η μέση τιμή (που συχνά λέγεται και μέσος όρος ή απλά «μέσος») χρησιμοποιείται ευρέως και είναι κατανοητή και από τον πολίτη μέσης μόρφωσης, η τυπική απόκλιση χρησιμοποιείται σπανιότερα και δεν είναι κατανοητή παρά μόνο από όσους κάνουν έρευνα πεδίου ή εμπειρική. Για το λόγο αυτό, και με την πεποίθηση ότι η τυπική απόκλιση δίνει μια διαφορετική και εξίσου σημαντική εικόνα ενός συνόλου δεδομένων, αφού φανερώνει το βαθμό ομοιογένειας ενός συνόλου ως προς μία μεταβλητή, εξηγώ σύντομα παρακάτω πώς υπολογίζεται, γιατί μόνο γνωρίζοντας πώς βγαίνει ένας στατιστικός δείκτης μπορεί κανείς να καταλάβει πλήρως πώς πρέπει να τον ερμηνεύει. Στη διάρκεια της πανεπιστημιακής μου σταδιοδρομίας έχω διαπιστώσει πως οι νέοι ερευνητές, δηλαδή οι μεταπτυχιακοί φοιτητές, που χρησιμοποιούν από την αρχή προγράμματα υπολογιστή για στατιστική επεξεργασία δεν αποκτούν ποτέ επίγνωση της σημασίας των δεικτών.

Η τυπική απόκλιση υπολογίζεται με τον τύπο:

$$SD = \sqrt{\frac{\sum(X - \bar{X})^2}{N}}$$

Στον τύπο αυτό το SD σημαίνει Standard Deviation (τυπική απόκλιση), το \sum σημαίνει σύνολο, άθροισμα, το X σημαίνει δεδομένο, τιμή, το \bar{X} σημαίνει μέση τιμή και το N (number) σημαίνει αριθμός δεδομένων ή τιμών ή υποκειμένων της έρευνας.

Όποιος επιχειρεί να υπολογίσει την τυπική απόκλιση «με το χέρι» πρώτα πρέπει να υψώσει στο τετράγωνο τη διαφορά κάθε βαθμού από τον αριθμητικό μέσο και μετά να κάνει την πρόσθεση.

Πρέπει, όμως, να γίνει εδώ σαφές ότι δεν προτείνεται ο υπολογισμός της τυπικής απόκλισης, που είναι βέβαια κουραστικός και χωρίς ενδιαφέρον, να γίνεται με το χέρι. Υπάρχουν μικρά και φτηνά «υπολογιστηράκια» (αυτά που έχουν συνήθως στο μοντέλο τους την ένδειξη «scientific calculator»), τα οποία «βγάζουν» αμέσως τον αριθμητικό μέσο και την τυπική απόκλιση. Σημασία έχει να γνωρίζει κανείς τον τύπο και να υπολογίσει μερικές φορές την τυπική απόκλιση χωρίς κομπιουτεράκι ή ηλεκτρονικό υπολογιστή για να κατανοήσει τι είναι πώς διαφοροποιείται με τη διαφοροποίηση των δεδομένων.

Αν δοθεί σε κάποιον η τιμή της τυπικής απόκλισης μιας κατανομής (για παράδειγμα $SD = 2,32$), από μόνη της δεν είναι ερμηνεύσιμη. Δεν επιτρέπει, δηλαδή, να καταλάβει κανείς πόσο σκόρπια είναι τα δεδομένα, άρα πόσο ανομοιογενής είναι η ομάδα υποκειμένων από την οποία προέκυψαν. Κι αυτό γιατί, αν, για παράδειγμα, η κλίμακα που χρησιμοποιήθηκε για τη μέτρηση της μεταβλητής είχε 10 μόνο βαθμίδες, τότε το 2,32 δείχνει αρκετά μεγάλη ανομοιογένεια, ενώ αν η κλίμακα είχε 100 βαθμίδες η ανομοιογένεια θα χαρακτηριζόταν ιδιαίτερα μικρή. Επομένως είναι απαραίτητο με κάθε

δείκτη διασποράς να γίνεται γνωστό και το συνολικό εύρος της κλίμακας που χρησιμοποιήθηκε. Στην περίπτωση μας η κλίμακα έχει 100 βαθμίδες.

Ακόμη και τότε, δεν υπάρχουν κοινώς αποδεκτά όρια για τον χαρακτηρισμό της διασποράς ως μικρής ή μεγάλης, πράγμα, αντίθετα, που γίνεται με τους δείκτες συνάφειας. Κατά συνέπεια η ερμηνεία του δείκτη απαιτεί και τον ad hoc καθορισμό συμβατικών ορίων, ανάλογα με την ερευνητική κατάσταση. Ελλείψει τέτοιων ορίων, η ερμηνεία των δεικτών διασποράς διευκολύνεται πολύ από τη δυνατότητα σύγκρισης των τιμών της τυπικής, π.χ., απόκλισης δύο ή περισσότερων ομάδων δεδομένων. και αυτό ακριβώς θα κάνουμε στο μεθεπόμενο κεφάλαιο.

2.2. Τρίτη ενότητα

Οι τρεις πίνακες που ακολουθούν περιλαμβάνουν τα δεδομένα της ίδιας εξεταστικής περιόδου, αλλά για την τρίτη ενότητα ή φάση της εξέτασης (κατανόηση προφορικού λόγου). Είναι χρήσιμο εδώ να διευκρινιστεί ότι τα ερωτήματα από 16 έως 20 στα οποία δεν εμφανίζεται το ποσοστό των ατόμων που επέλεξαν κάθε δυνατή απάντηση είναι, ακριβώς, ερωτήματα που δεν παρέχουν στους εξεταζόμενους δυνατές απαντήσεις. Δηλαδή τους ζητείται να γράψουν κάτι κι όχι να επιλέξουν μαυρίζοντας ένα κουτάκι.

ΜΑΙΟΣ 2007 B1

	e.i.*100	d.i.*100	A	B	C	D	E	x		
1	58	66	34	58	7			0		
2	23	53	27	50	23			0		
3	69	59	27	69	3			0	1	
4	66	59	66	15	19			0	52,0	
5	44	36	52	4	44			0	54,6	
6	74	39	18	8	74			0		
7	29	36	62	29	8			0		
8	28	46	14	58	28			0	2	
9	50	45	29	50	21			0	54,0	
10	89	20	89	6	5			0	37,2	
11	39	44	17	39	42			0		
12	12	24	54	34	12			0		
13	29	34	29	15	55			0	3	
14	65	40	17	65	18			0	42,2	
15	66	32	66	18	15			0	34,8	
16	47	64								
17	4	4								
18	9	15							4	
19	12	30							24,6	
20	51	59							34,4	
	43,20	40,25	average							
	23,75	16,42	st.dev.							

ΜΑΙΟΣ 2007 Β2

	e.i.*100	d.i.*100	A	B	C	D	E	x	
1	47	44	53	47				0	
2	84	33	84	16				0	
3	75	39	25	75				0	
4	51	35	51	49				0	
5	93	19	7	93				0	
6	88	14	88	12				0	1
7	51	33	51	49				0	71,5
8	83	11	83	17				0	28,5
9	78	37	78	5	17			0	
10	78	42	6	78	16			0	
11	63	47	24	63	12			0	
12	77	43	19	4	77			0	2
13	65	49	7	28	65			0	73,2
14	78	12	5	17	78			0	38,3
15	93	17	93	1	6			0	
16	78	49	16	6	78			0	
17	95	10	4	95	1			0	
18	92	14	4	5	92			0	3
19	73	41	73	10	16			0	80,5
20	52	48	35	13	52			0	29,8
21	73	43							
22	23	42							
23	33	50							4
24	14	29							29,4
25	4	12							35,2
	65,64	32,52	average						
	24,86	14,01	st.dev.						

ΜΑΙΟΣ 2007 Γ1

	e.i.*100	d.i.*100	A	B	C	D	E	x	
1	96	10	3	96	1			0	
2	90	28	90	4	6			0	
3	91	16	91	8	1			0	
4	70	42	7	23	70			0	
5	92	16	4	92	4			0	
6	81	35	81	8	11			0	
7	49	61	40	10	49			0	1
8	50	36	50	12	38			0	76,8
9	72	54	20	7	72			0	33,1
10	87	24	87	13				0	
11	97	8	3	97				0	
12	82	25	82	18				0	
13	57	51	43	57				0	2
14	83	21	83	17				0	76,2
15	51	39	51	48				0	28,0
16	69	29	69	31				0	
17	90	20	10	90				0	
18	74	21	74	26				0	
19	35	32	65	35				0	

20	42	36	58	42				0	
21	18	40							
22	33	52							
23	11	24							3
24	15	40							39,7
25	10	30							32,4
	61,80	31,60	average						
	28,11	13,51	st.dev.						

3. ΔΕΙΓΜΑ ΔΕΔΟΜΕΝΩΝ ΑΠΟ ΘΕΜΑΤΑ ΑΛΛΩΝ ΓΛΩΣΣΩΝ

Παρόλο ότι δεν προβλέπετο από το Τεχνικό Δελτίο του Υποέργου, ως επιστημονικώς υπεύθυνος αποφάσισα ό,τι είδους έλεγχο έκανα για τα όργανα μέτρησης της ιταλικής να κάνω και για τις υπόλοιπες γλώσσες, για να είναι δυνατή η σύγκριση και η εξαγωγή συμπερασμάτων. Για εξοικονόμηση χώρου θα δοθούν παρακάτω οι πίνακες των δεικτών μόνο για το επίπεδο B1 σε καθεμία από τις γλώσσες και μόνο για την περίοδο Μαΐου 2007. Δεδομένα, ωστόσο, διαθέτω για όλες τις εξεταστικές περιόδους, για όλες τις γλώσσες, για όλα τα επίπεδα και για όλες τις ενότητες της εξέτασης.

3.1. Αγγλική γλώσσα

Τα δεδομένα για την αγγλική γλώσσα προέκυψαν από δείγμα 29.039 ατόμων, συνολικά. Η κατανομή των υποκειμένων εμφανίζεται στη σελίδα 8.

ΜΑΪΟΣ 2007 Β1. ΕΝΟΤΗΤΑ 1

	e.i.*100	d.i.*100	A	B	C	D	E	F	G	H	x	
1	81	40	7	81	4	1	3	3			1	
2	89	19	7	1	1	0	89	0			1	
3	72	43	72	2	17	2	3	2			1	1
4	84	37	1	10	1	1	1	84			1	84,0
5	94	19	2	1	1	94	1	1			1	31,6
6	89	25	89	4	6						1	
7	73	22	2	73	24						1	
8	85	25	85	7	7						1	
9	69	46	18	69	11						1	
10	69	43	22	7	69						2	
11	80	33	4	15	80						1	
12	82	27	8	82	9						1	
13	37	53	54	37	8						1	
14	38	39	38	12	49						2	
15	39	22	39	36	23						2	
16	40	25	36	40	22						2	2
17	47	11	24	26	47						2	62,2
18	61	42	8	29	61						2	31,8
19	88	31	2	1	2	2	2	1	0	88	2	
20	38	59	38	4	10	6	19	16	3	2	3	
21	82	39	3	2	3	2	2	3	82	1	2	
22	64	58	9	3	64	7	6	7	2	1	2	3
23	40	56	12	40	6	8	11	16	3	1	4	56,2
24	25	38	14	12	6	15	25	21	2	1	5	46,8
25	50	54	9	6	50	13	3	2	11	2	4	
26	18	40	12	7	20	15	2	11	18	8	7	
27	41	51	10	5	4	41	11	6	14	4	5	
28	46	75	8	7	11	5	2	7	8	46	6	
29	30	52	30	10	4	7	10	13	13	7	6	
30	31	59	12	10	7	6	6	31	14	9	6	4
31	49	68	8	49	2	4	5	16	6	6	5	39,9
32	54	55	8	5	3	6	54	5	6	5	7	56,8
33	40	15	34	40	23						4	
34	55	38	55	26	15						4	
35	64	51	20	64	12						3	
36	71	43	8	14	71						7	
37	74	34	74	13	9						3	
38	40	27	23	31	40						6	5
39	57	28	57	19	20						3	55,4
40	42	9	42	32	22						4	30,6
41	70	38										
42	51	65										
43	27	46										6
44	28	64										42,0
45	34	55										53,6
46	8	19										
47	56	60										
48	35	65										7
49	20	41										29,0
50	26	56										48,2
	53,66	41,20	average									
	21,97	15,98	st.dev.									

ΜΑΪΟΣ 2007 Β1. ΕΝΟΤΗΤΑ 3

	e.i.*100	d.i.*100	A	B	C	D	E	F	x		
1	34	46	14	34	9	6	26	9	1		
2	59	62	10	12	8	59	7	3	2		
3	23	34	16	12	31	8	23	7	2	1	
4	51	63	12	10	7	7	11	51	1	40,2	
5	34	51	34	15	13	10	18	9	2	130,0	
6	47	58	44	9	47				1		
7	39	60	57	39	3				0		
8	38	41	38	50	11				1	2	
9	61	50	23	15	61				1	46,0	
10	45	48	18	45	37				1	148,8	
11	86	29	86	6	8				0		
12	78	32	7	78	14				1		
13	79	40	79	5	15				1	3	
14	36	24	14	36	49				1	65,4	
15	48	54	18	32	48				1	196,4	
16	16	37									
17	35	41									
18	15	28								4	
19	11	22								17,0	
20	8	20								58,4	
	42,15	133,40	average								
	21,82	63,04	st.dev.								

3.2. Γαλλική γλώσσα

Τα δεδομένα για τη γαλλική γλώσσα προέκυψαν από δείγμα 4.411 ατόμων, συνολικά. Η κατανομή των υποκειμένων εμφανίζεται και πάλι στη σελίδα 8. Η μόνη περίπτωση κατά την οποία ο αριθμός των ατόμων είναι πολύ μικρός, τόσος που να μην μπορεί να εξαχθεί ασφαλές αποτέλεσμα, είναι το επίπεδο Γ1 περιόδου Νοεμβρίου 2007.

ΜΑΪΟΣ 2007 Β1. ΕΝΟΤΗΤΑ 1

	e.i.*100	d.i.*100	A	B	C	D	E	F	G	H	x	
1	85	24	85	0	14						0	
2	97	3	1	97	1						0	
3	92	10	6	92	1						1	1
4	90	22	7	0	90						2	89,2
5	82	14	16	82	1						1	14,6
6	97	7	2	0	97	0	0				0	
7	83	29	6	0	0	83	9				0	

8	90	19	90	1	1	2	6				0	2
9	85	24	2	0	0	13	85				0	90,8
10	99	3	0	99	0	0	0				0	16,4
11	47	22	45	47	7						0	
12	59	45	59	33	7						0	
13	52	40	52	34	14						0	3
14	69	41	24	69	8						0	62,4
15	85	31	11	5	85						0	35,8
16	87	22	0	1	10	87	1				0	
17	56	55	15	2	56	6	21				0	
18	73	50	73	1	13	3	9				0	4
19	68	48	9	0	19	3	68				1	75,8
20	95	9	3	95	1	0	0				0	36,8
21	70	31	70	17	12						1	
22	86	29	4	10	86						0	
23	79	29	15	79	6						0	5
24	65	41	7	65	28						0	64,6
25	23	52	71	23	6						0	36,4
26	77	29	77	12	11						0	
27	50	34	2	48	50						0	
28	62	45	23	62	15						0	6
29	74	38	5	21	74						0	66,6
30	70	45	70	18	12						0	38,2
31	35	12	48	35	18						0	
32	69	33	5	69	26						0	
33	36	36	7	36	56						0	7
34	89	26	7	4	89						0	58,2
35	62	48	62	32	6						0	31,0
36	44	19	50	44	6						0	
37	59	57	30	59	11						0	
38	75	38	5	20	75						0	8
39	69	34	69	23	8						0	63,2
40	69	41	69	14	17						0	37,8
41	27	40										
42	28	57										
43	28	24										9
44	56	52										38,0
45	51	43										43,2
46	97	12										
47	87	24										
48	48	38										10
49	94	19										84,4
50	96	5										19,6
	69,32	30,98	average									
	21,01	14,70	st.dev.									

ΜΑΪΟΣ 2007 Β1. ΕΝΟΤΗΤΑ 3

	e.i.*100	d.i.*100	A	B	C	D	E	x		
1	33	52	12	14	33	5	36	0		
2	53	62	53	7	16	14	9	0		
3	40	40	12	5	20	25	40	0	1	
4	67	60	10	67	8	11	3	0	47,8	
5	46	71	14	9	21	46	10	0	57,0	
6	63	48	28	63	9			0		
7	43	55	43	36	20			0		
8	90	26	5	5	90			0	2	
9	82	12	5	82	13			0	69,4	
10	69	55	69	18	13			0	39,2	
11	70	72	2	10	13	5	70	0		
12	93	21	93	1	2	1	2	0		
13	78	55	0	78	9	4	8	0	3	
14	87	33	1	4	4	87	5	0	80,0	
15	72	69	4	8	72	3	13	0	50,0	
16	78	34								
17	55	55								
18	67	52							4	
19	77	33							62,0	
20	33	55							45,8	
	64,80	48,00	average							
	18,07	16,40	st.dev.							

3.3. Γερμανική γλώσσα

Τέλος, για τη γερμανική γλώσσα είχαμε ένα δείγμα μεγέθους 6.546 ατόμων, που κατανέμονταν γεωγραφικά έτσι όπως φαίνεται στη σελίδα 8.

Τα δεδομένα Μαΐου 2007 και μόνο για το Β1 είναι τα εξής:

ΜΑΪΟΣ 2007 Β1. ΕΝΟΤΗΤΑ 1

	e.i.*100	d.i.*100	A	B	C	D	E	F	G	H	x	
1	91	21	1	91	0	0	0	5	2		0	
2	96	8	96	1	0	0	1	1	1		0	
3	97	7	0	0	0	97	0	0	1		1	
4	97	8	0	0	97	0	0	3	0		0	
5	91	26	1	3	0	0	91	2	3		0	1
6	79	51	1	2	1	1	3	14	79		0	89,4
7	75	57	1	2	2	0	5	75	14		0	25,4
8	77	49	77	22							1	
9	82	45	18	82							0	
10	69	30	69	31							0	
11	59	44	41	59							0	
12	88	27	12	88							0	
13	54	41	54	45							1	2
14	53	16	47	53							0	67,3

15	56	37	43	56						0	36,1
16	86	27	4	9	86					0	
17	98	9	98	1	1					0	
18	73	42	10	18	73					0	
19	83	36	10	83	7					0	3
20	71	53	8	21	71					0	80,7
21	73	46	9	73	18					0	35,5
22	53	71	8	16	11	13	53			0	
23	62	53	62	5	8	6	18			0	
24	66	62	10	66	5	9	9			0	4
25	50	63	19	8	11	50	12			1	59,2
26	65	57	2	3	65	22	8			0	61,2
27	75	45	16	8	75					1	
28	61	62	16	22	61					1	
29	28	23	27	45	28					0	
30	7	-11	71	22	7					0	
31	71	55	9	20	71					0	5
32	57	45	57	30	13					0	52,3
33	67	55	67	16	17					0	39,1
34	97	5	97	2	1	0	0	0	0	0	
35	95	12	0	1	95	1	1	0	1	0	
36	81	47	1	5	4	81	4	1	3	0	
37	82	41	1	82	0	13	0	2	1	1	
38	93	16	0	4	0	1	1	93	0	1	6
39	73	40	0	0	0	0	25	1	73	0	84,3
40	69	47	0	3	0	4	69	1	20	1	29,7
41	65	31									
42	44	74									
43	36	52									
44	61	50									
45	52	48									
46	53	55									
47	77	55									
48	17	44									7
49	47	48									50,8
50	56	84									54,1
	68,16	40,18	average								
	20,33	19,72	st.dev.								

ΜΑΪΟΣ 2007 Β1. ΕΝΟΤΗΤΑ 3

	e.i.*100	d.i.*100	A	B	C	D	E	x	
1	83	41	11	83	1	0	4	0	
2	91	19	3	2	91	1	3	0	
3	75	53	75	6	3	3	12	1	1
4	92	24	2	3	1	92	3	0	84,0
5	79	53	9	7	3	2	79	0	38,0
6	70	41	70	30				0	
7	48	25	52	48				0	
8	28	27	72	28				0	
9	72	47	72	28				0	
10	56	16	56	43				0	
11	49	36	51	49				0	

12	74	41	74	26				1	
13	74	48	25	74				0	2
14	63	52	63	37				0	62,5
15	91	9	91	8				0	34,2
16	79	37							
17	92	22							
18	82	34							3
19	70	44							73,2
20	43	55							38,4
	70,55	36,20	average						
	17,38	13,35	st.dev.						

4. ΣΥΓΚΡΙΤΙΚΗ ΑΝΑΛΥΣΗ

Ενώ οι ειδικοί διαφοροποιούνται μεταξύ τους ως προς τον τρόπο υπολογισμού των παραπάνω δύο δεικτών, όλοι συμφωνούν, όπως ήδη αναφέραμε αλλού, πως ο πρώτος (δείκτης ευκολίας) επιτρέπει να το περιγράψουμε καθώς και να πούμε πόσο κατάλληλο ήταν για τους συγκεκριμένους υποψήφιους ενώ ο δεύτερος (δείκτης διάκρισης) φανερώνει την ποιότητα των θεμάτων εξέτασης, ως προς την ικανότητά τους να εντοπίζουν αυτούς που είναι σε θέση να επικοινωνήσουν με την ξένη γλώσσα λιγότερο ή περισσότερο καλά.

Στο κεφάλαιο αυτό συγκρίνονται οι δείκτες των θεμάτων ιταλικής με αυτούς των άλλων γλωσσών.

4.1. Πρώτη ενότητα

Στον πίνακα της επόμενης σελίδας φαίνονται, συγκεντρωτικά, οι μέσες τιμές και οι τυπικές αποκλίσεις από τη χρήση των οργάνων μέτρησης που ετοίμασαν οι επιστημονικές ομάδες όλων των γλωσσών και για τις τρεις εξεταστικές περιόδους για τα επίπεδα A1 και A2 που εξετάζονται με ενιαίο τεστ. Στη συνέχεια θα καταγραφούν τα δεδομένα για το B1, το B2 και το Γ1 (τέσσερις συνολικά πίνακες), δηλαδή όλα τα επίπεδα στα οποία η Κεντρική Εξεταστική Επιτροπή του ΚΠΓ έκρινε σκόπιμο και εφικτό να διεξαχθούν εξετάσεις πανελλαδικά. Όργανα μέτρησης για το επίπεδο Γ2 στην ιταλική γλώσσα δεν ετοιμάστηκαν και ούτε είχαν προβλεφτεί στο ΤΔΥ. Ο λόγος είναι πως η ΚΕΕ

έχει αποφασίσει, από την αρχή λειτουργίας του συστήματος, ότι όσο δεν τροποποιείται ο θεσμός της «επάρκειας» δεν θα γίνουν εξετάσεις επιπέδου Γ2.

Επίπεδα A1 και A2 (ενιαίο τεστ). Μάιος 2008

Γλώσσα	Ευκολία ερωτημάτων (καταλληλότητα του τεστ)		Διακριτική ικανότητα ερωτημάτων (ποιότητα του τεστ)	
	Μέση τιμή (κεντρική τάση)	Τυπική απόκλιση (ομοιογένεια)	Μέση τιμή (κεντρική τάση)	Τυπική απόκλιση (ομοιογένεια)
Αγγλική	70,36	17,64	40,34	18,70
Γαλλική	69,48	22,84	36,24	19,58
Γερμανική	80,78	17,25	33,54	24,15
Ιταλική	76,50	23,28	26,64	19,87

Από τα παραπάνω δεδομένα προκύπτει πως το όργανο μέτρησης για την ιταλική γλώσσα στο συγκεκριμένο επίπεδο και στη συγκεκριμένη εξεταστική περίοδο είχε τα εξής χαρακτηριστικά:

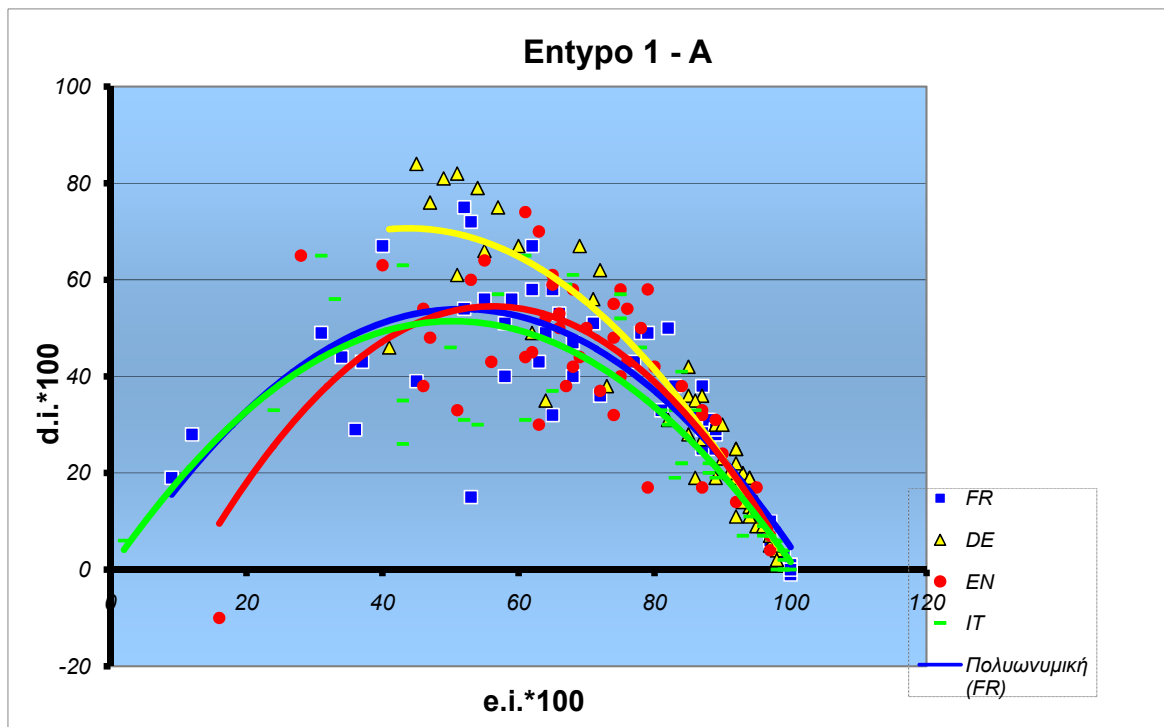
- α) Ως προς την **καταλληλότητα** για το κοινό που εξετάστηκε, ήταν πολύ καλό, αφού ο μέσος δείκτης ευκολίας ήταν 76,50. Επειδή πρόκειται για τα χαμηλότερα επίπεδα γλωσσομάθειας, θεωρώ, συμβατικά, πως η ευκολία πρέπει να είναι μεταξύ 60 και 80. Μόνο το όργανο για τη γερμανική υπερβαίνει λίγο το συμβατικό όριο που εγώ όρισα. Με άλλα λόγια ήταν υπερβολικά εύκολο.
- β) Ως προς την **ποιότητα**, ήταν το λιγότερο καλό από όλα τα άλλα. Θεωρείται, ωστόσο, καλής ποιότητας, έστω και οριακά (26,64, όταν το όριο μεταξύ μέτριας και καλής ποιότητας είναι το 26,00), σύμφωνα με τη βιβλιογραφία (βλέπε και σελίδα 15).

Όργανο κακής ποιότητας	Όργανο μέτριας ποιότητας	Όργανο καλής ποιότητας
-1,00 έως 0,00	+0,01 έως +0,25	+0,26 έως +1,00

- γ) Ως προς την **ομοιογένεια της ευκολίας**, που θα μπορούσε να θεωρηθεί ως μία διάσταση της ποιότητας, το όργανο αποδείχτηκε καλύτερο από όλα τα άλλα, με την υψηλότερη τυπική απόκλιση της ευκολίας στο 23,28. Η τυπική απόκλιση του δείκτη ευκολίας πρέπει να είναι υψηλή (πάνω από 20,00) όταν ένα τεστ είναι διαβαθμισμένο και να είναι μικρή (κάτω από 20,00) όταν το τεστ έχει ετοιμαστεί για ένα και μόνο επίπεδο.
- δ) Τέλος, ως προς την **ομοιογένεια της διακριτικής ικανότητας**, το τεστ ιταλικής ήταν καλό, όπως άλλωστε και τα τεστ αγγλική και γαλλικής. Μόνο το τεστ γερμανικής υπερβαίνει, και μάλιστα κατά πολύ το συμβατικό όριο του 20,00 που εμείς θέτουμε, αδιάφορα με το αν το τεστ είναι διαβαθμισμένο ή όχι. Με άλλα λόγια, όταν ένα τεστ είναι για περισσότερα του ενός επίπεδα πρέπει να περιλαμβάνει δοκιμασίες τόσο εύκολες όσο και δύσκολες, επομένως να έχει τυπική απόκλιση ευκολίας μεγαλύτερη του 20,00. Αναφορικά, όμως, με τη διακριτική ικανότητα αυτό δεν ισχύει επειδή είναι επιθυμητό όλες οι δοκιμασίες να έχουν λίγο ως πολύ την ίδια ποιότητα είτε πρόκειται για διαβαθμισμένο τεστ είτε όχι.

Με τον ίδιο τρόπο με τον οποίο παρουσιάστηκαν και σχολιάστηκαν τα δεδομένα για τα επίπεδα A1 και A1 μίας εξεταστικής περιόδου θα μπορούσαν να εξεταστούν και για όλα

τα άλλα επίπεδα των τριών περιόδων, πράγμα που θα απαιτούσε πολλές επιπλέον σελίδες σε αυτό το Παραδοτέο. Για λόγους οικονομίας θα περιοριστούμε σε τρεις ακόμη ανάλογες περιπτώσεις. Πριν, όμως, από αυτό θα δούμε παρακάτω ένα στικτό διάγραμμα (scatterplot) όπου απεικονίζονται ταυτόχρονα οι τιμές ευκολίας και διάκρισης όλων των ερωτημάτων που υπήρχαν στα τεστ των τεσσάρων γλωσσών. Παρόμοια διαγράμματα ετοίμασα, ως επιστημονικώς υπεύθυνος του Υποέργου, για όλα τα επίπεδα και όλες τις ενότητες, αλλά δεν θα τα παρουσιάσουμε και σχολιάσουμε στο κυρίως κείμενο. Περιλαμβάνονται, ωστόσο, όλα στο παράρτημα, επειδή έχω την πεποίθηση ότι τα στικτά αυτά διαγράμματα αποτελούν την «ακτινογραφία» των οργάνων που ετοιμάστηκαν στο πλαίσιο αυτού του προγράμματος ΕΠΕΑΕΚ: ένας γνώστης της «ανάλυσης ερωτημάτων» μπορεί από τα στικτά αυτά διαγράμματα και μόνο να βγάλει συμπεράσματα για την ποιότητα των οργάνων σε όλες τις γλώσσες και ειδικότερα, αυτό που μας ενδιαφέρει εδώ, της ιταλικής, σε σχέση και με τις άλλες.



Μερικά σχόλια που μπορεί να κάνει κανείς είναι:

- α) Η κορυφή της ομαλής καμπύλης που απεικονίζει την πολυωνυμική γραμμή τάσης των δεικτών για την αγγλική γλώσσα (κόκκινο χρώμα) είναι πιο πάνω από όλες τις άλλες. Αυτό δείχνει, με εποπτικό τρόπο πως η αγγλική έβγαλε το μεγαλύτερο δείκτη διάκρισης, δηλαδή είχε το καλύτερο τεστ σε αυτή την περίοδο, σε αυτό το επίπεδο και σε αυτή την ενότητα της εξέτασης.
- β) Η γραμμή τάσης της γερμανικής γλώσσας είναι διαφορετική από τις άλλες και αυτό οφείλεται στον υπέρμετρα υψηλό δείκτη ευκολίας και, γενικά, σε κάποια ανισορροπία που παρουσίασε το όργανο μέτρησης.
- γ) Κάτω από το 0, δηλαδή ερωτήματα κακής ποιότητας, φαίνεται πως υπήρχαν μόνο δύο: ένα της αγγλικής (κόκκινη κουκίδα) και ένα της γαλλικής (μπλε τετράγωνο).

δ) Η γραμμή τάσης της ιταλικής είναι η περισσότερο συμμετρική. Επομένως, η καμπύλη έχει την κορυφή της χαμηλότερα από όλες τις άλλες (δηλαδή ήταν τεστ με τη μικρότερη ικανότητα διάκρισης) είχε όμως μια ισορροπία, ως προς τη δυσκολία (δηλαδή απευθυνόταν περίπου με το ίδιο ποσοστό ερωτημάτων τόσο σε άτομα με ετοιμασία για το Α1 όσο και σε άτομα που είχαν ετοιμαστεί για το Α2.

Ακολουθούν ένας πίνακες δεδομένων για τα άλλα επίπεδα, από διαφορετικές περιόδους.

Επίπεδο Β1. Μάιος 2007

Γλώσσα	Ευκολία ερωτημάτων		Διακριτική ικανότητα ερωτημάτων	
	Μέση τιμή	Τυπική απόκλιση	Μέση τιμή	Τυπική απόκλιση
Αγγλική	53,66	21,97	41,20	15,98
Γαλλική	69,32	21,01	30,98	14,70
Γερμανική	68,16	20,33	40,18	19,72
Ιταλική	65,30	22,44	39,36	16,62

Επίπεδο Β2. Νοέμβριος 2007

Γλώσσα	Ευκολία ερωτημάτων		Διακριτική ικανότητα ερωτημάτων	
	Μέση τιμή	Τυπική απόκλιση	Μέση τιμή	Τυπική απόκλιση
Αγγλική	60,80	24,45	37,56	14,06
Γαλλική	55,44	26,50	30,04	13,50
Γερμανική	68,58	17,54	38,96	17,36
Ιταλική	51,96	21,00	35,00	14,33

Επίπεδο Γ1. Μάιος 2008

Γλώσσα	Ευκολία ερωτημάτων		Διακριτική ικανότητα ερωτημάτων	
	Μέση τιμή	Τυπική απόκλιση	Μέση τιμή	Τυπική απόκλιση
Αγγλική	61,62	19,53	34,43	12,96
Γαλλική	60,63	25,80	29,53	14,48
Γερμανική	52,60	19,27	43,77	19,32
Ιταλική	61,05	21,74	38,55	15,01

Ο σχολιασμός και η ερμηνεία των δεδομένων στη σελίδα 26 δίνουν στον αναγνώστη αυτής της Αναφοράς την κλειδα ερμηνείας. Θα περιοριστούμε εδώ σε λίγα σχόλια πάνω στα δεδομένα των τριών παραπάνω πινάκων:

α) Το Μάιο του 2007, στην πρώτη ενότητα (κατανόηση γραπτού λόγου και γλωσσική επίγνωση) το καλύτερο τεστ ήταν αυτό της αγγλικής, ακολουθούμενο, με μικρή διαφορά από το τεστ γερμανικής και ιταλικής. Στην περίοδο αυτή το λιγότερο καλό τεστ ήταν στη γαλλική γλώσσα, που, ωστόσο, εξακουθούσε να θεωρείται καλό, αφού ο μέσος δείκτης διάκρισης που έβγαλε ήταν μεγαλύτερος από το 26,00.

- β) Το Νοέμβριο του 2007 επαναλαμβάνεται η ίδια εικόνα, δηλαδή το τεστ αγγλικής ήταν το καλύτερο, ακολουθούμενο από το τεστ γερμανικής και ιταλικής.
- γ) Το Μάιο του 2008, τέλος, έχουμε το τεστ γερμανικής να δίνει την καλύτερη εικόνα, ακολουθούμενο από το τεστ ιταλικής, μετά αγγλικής και τέλος, για τρίτη φορά, λιγότερο καλό φαίνεται το τεστ γαλλικής, αλλά αρκετά πάνω από το όριο που χωρίζει ένα μέτριο τεστ από ένα καλό τεστ.

4.2. Τρίτη ενότητα

Στο υποκεφάλαιο αυτό θα δοθούν μερικά δεδομένα από τα τεστ της τρίτης ενότητα, που έχει ως στόχο τον έλεγχο της ικανότητας κατανόησης προφορικού λόγου, στον οποίο οι εξεταζόμενοι εκτίθενται μέσω μεγαφωνικών εγκαταστάσεων στις αίθουσες και χρήση αυθεντικών, κυρίως, κειμένων.

Επίπεδα A1 και A2 (ενιαίο τεστ). Μάιος 2008

Γλώσσα	Ευκολία ερωτημάτων		Διακριτική ικανότητα ερωτημάτων	
	Μέση τιμή	Τυπική απόκλιση	Μέση τιμή	Τυπική απόκλιση
Αγγλική	66,20	20,85	46,10	23,57
Γαλλική	72,60	15,27	48,40	18,72
Γερμανική	75,00	15,44	46,15	16,90
Ιταλική	64,00	30,22	42,60	27,08

Επίπεδο B1. Μάιος 2007

Γλώσσα	Ευκολία ερωτημάτων		Διακριτική ικανότητα ερωτημάτων	
	Μέση τιμή	Τυπική απόκλιση	Μέση τιμή	Τυπική απόκλιση
Αγγλική	42,15	21,82	42,00	13,29
Γαλλική	64,80	18,07	48,00	16,40
Γερμανική	70,55	17,38	36,20	13,35
Ιταλική	43,20	23,75	40,25	16,42

Επίπεδο B2. Νοέμβριος 2007

Γλώσσα	Ευκολία ερωτημάτων		Διακριτική ικανότητα ερωτημάτων	
	Μέση τιμή	Τυπική απόκλιση	Μέση τιμή	Τυπική απόκλιση
Αγγλική	51,60	20,44	37,85	17,48
Γαλλική	65,80	15,59	40,85	17,09
Γερμανική	74,35	12,16	37,95	15,11
Ιταλική	42,60	18,80	30,20	16,57

Επίπεδο Γ1. Μάιος 2008

Γλώσσα	Ευκολία ερωτημάτων		Διακριτική ικανότητα ερωτημάτων	
	Μέση τιμή	Τυπική απόκλιση	Μέση τιμή	Τυπική απόκλιση
Αγγλική	53,08	21,17	36,96	10,90
Γαλλική	62,76	22,88	30,56	12,90
Γερμανική	68,80	17,39	40,16	21,19
Ιταλική	62,40	25,68	37,56	13,62

Ο σχολιασμός των παραπάνω δεδομένων διευκολύνεται από το φόντο ορισμένων φατνίων. Τα φατνία με ανοικτό πράσινο φόντο περιλαμβάνουν δείκτη για τον οποίο θα έπρεπε να γίνει θετικό σχόλιο, ενώ τα φατνία με πορτοκαλί φόντο θα επέτρεπαν αρνητικό σχόλιο. Χάρη στην κλείδα ερμηνείας, είναι, νομίζω επιτρεπτό να αφεθεί στον αναγνώστη η εξαγωγή συμπερασμάτων αναφορικά με τα μεμονωμένα στοιχεία, για να περάσουμε στην εξαγωγή συμπερασμάτων αναφορικά με την ποιότητα των θεμάτων εξέτασης της ιταλικής.

5. ΣΥΜΠΕΡΑΣΜΑΤΑ

Τα δεδομένα που προέκυψαν από την υλοποίηση του παρόντος Πακέτου Εργασίας ήταν, φυσικά, πολύ περισσότερα από αυτά που δόθηκαν παραπάνω. Τα συμπεράσματά μας λαμβάνουν υπόψη το σύνολο των δεδομένων και όχι το μικρό δείγμα δεδομένων που περιλαμβάνονται στα προηγούμενα κεφάλαια.

Η διαδικασία που ακουθήθηκε στο πλαίσιο του Πακέτου Εργασίας Νο 7 για την αποτίμηση της ποιότητας των οργάνων μέτρησης που χρησιμοποιήθηκαν, με βασικό εργαλείο ελέγχου την ανάλυση των ερωτημάτων, μας επιτρέπει να βγάλουμε τα παρακάτω βασικά συμπεράσματα:

- α) Τα όργανα μέτρησης της ικανότητας επικοινωνίας μέσω της ιταλικής γλώσσας, ιδιαίτερα σε επίπεδο πρόσληψης γραπτών και ακουστικών μηνυμάτων, φάνηκε να είναι όλα καλής ποιότητας, αφού οι δείκτες ποιότητας υπερβαίνουν πάντα την κριτική τιμή που ορίζεται στη διεθνή βιβλιογραφία.
- β) Συγκριτικά με όλες τις άλλες γλώσσες, ωστόσο, τα όργανα μέτρησης της ιταλικής παρουσιάζουν λιγότερο καλή εικόνα, όπως φαίνεται από τον πίνακα της επόμενης σελίδας, με 35,87 μέσο δείκτη διάκρισης.
- γ) Η ομοιογένεια της ποιότητας των οργάνων ιταλικής είναι ικανοποιητική. Ως προς αυτό το κριτήριο η γαλλική παρουσιάζει την λιγότερο καλή εικόνα.

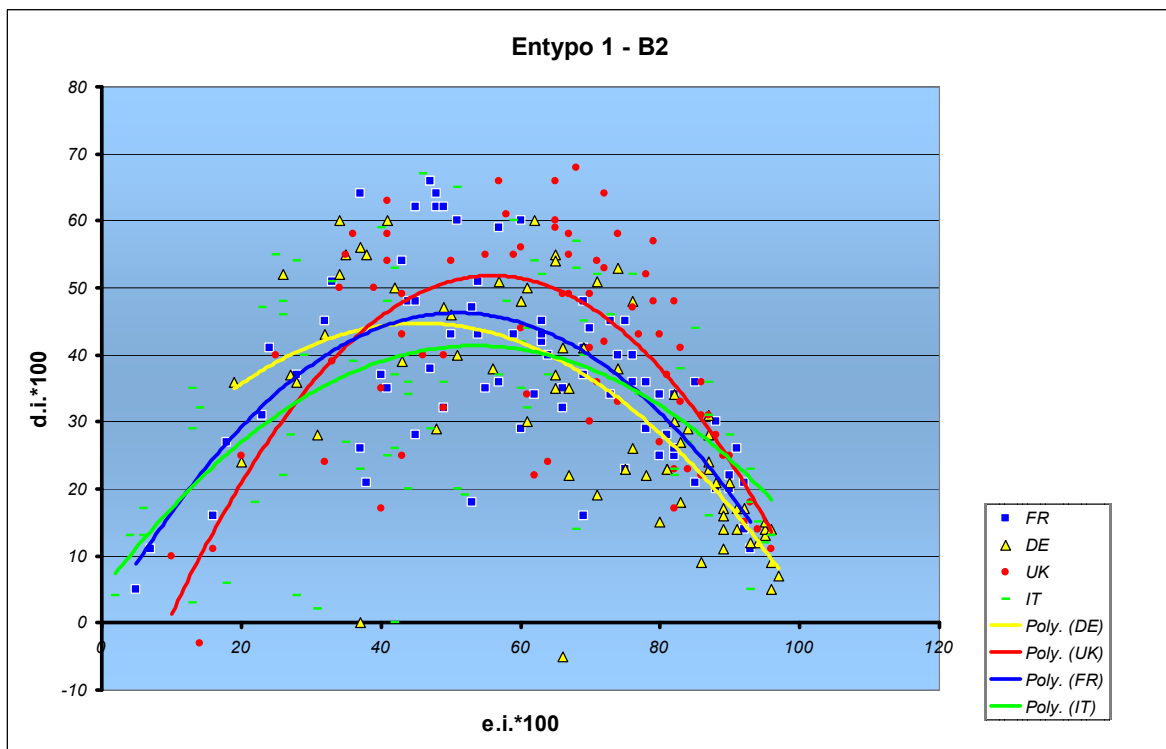
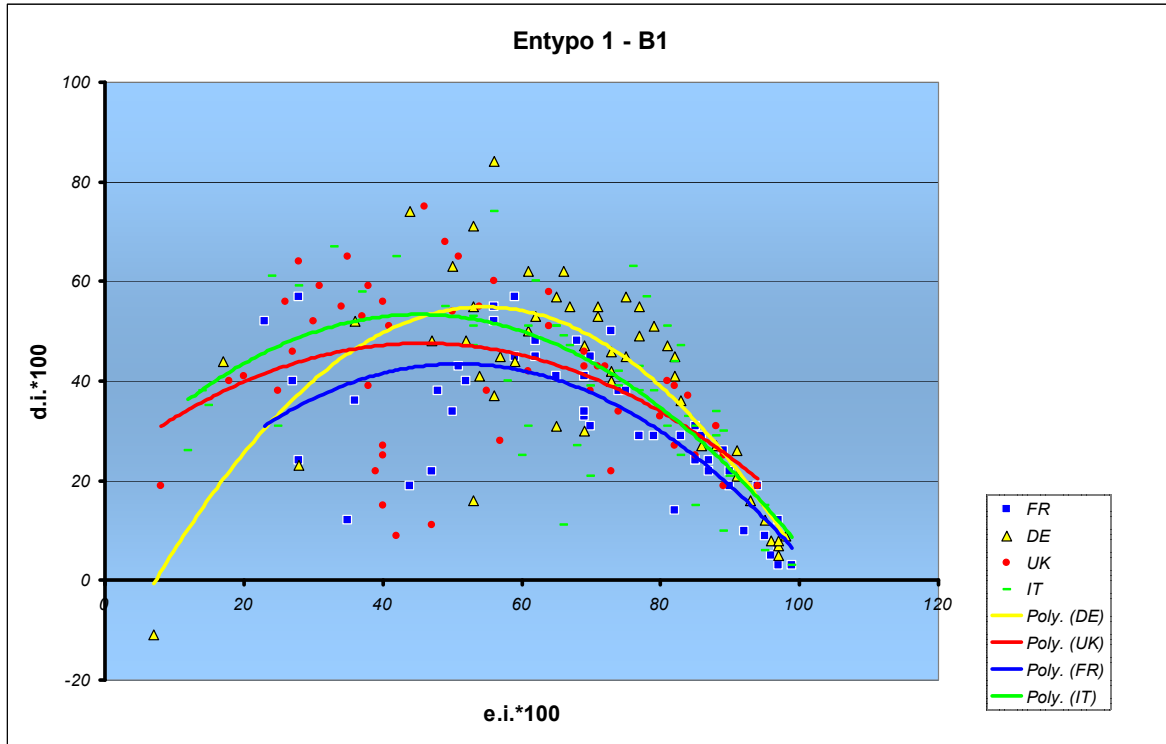
- δ) Το όργανο μέτρησης με την καλύτερη ποιότητα σε αυτές τις τρεις περιόδους είναι αυτό για το Β1 το Νοέμβριο του 2007 για την τρίτη ενότητα, στην αγγλική γλώσσα.
- ε) Το μόνο όργανο που έπεσε κάτω από το όριο μέτριας και καλής ποιότητας είναι αυτό που χρησιμοποιήθηκε για το επίπεδο Γ1 το Νοέμβριο του 2007 για την τρίτη ενότητα, στη γαλλική γλώσσα.
- στ) Μία αποτίμηση, τέλος, της ποιότητας όλων των οργάνων μέτρησης που ετοιμάστηκαν στο πλαίσιο των δύο Έργων που ανέλαβαν το ΕΚΠΑ και το ΑΠΘ φανερώνει υπεροχή των οργάνων που ετοίμασε η επιστημονική ομάδα αγγλικής του ΕΚΠΑ.

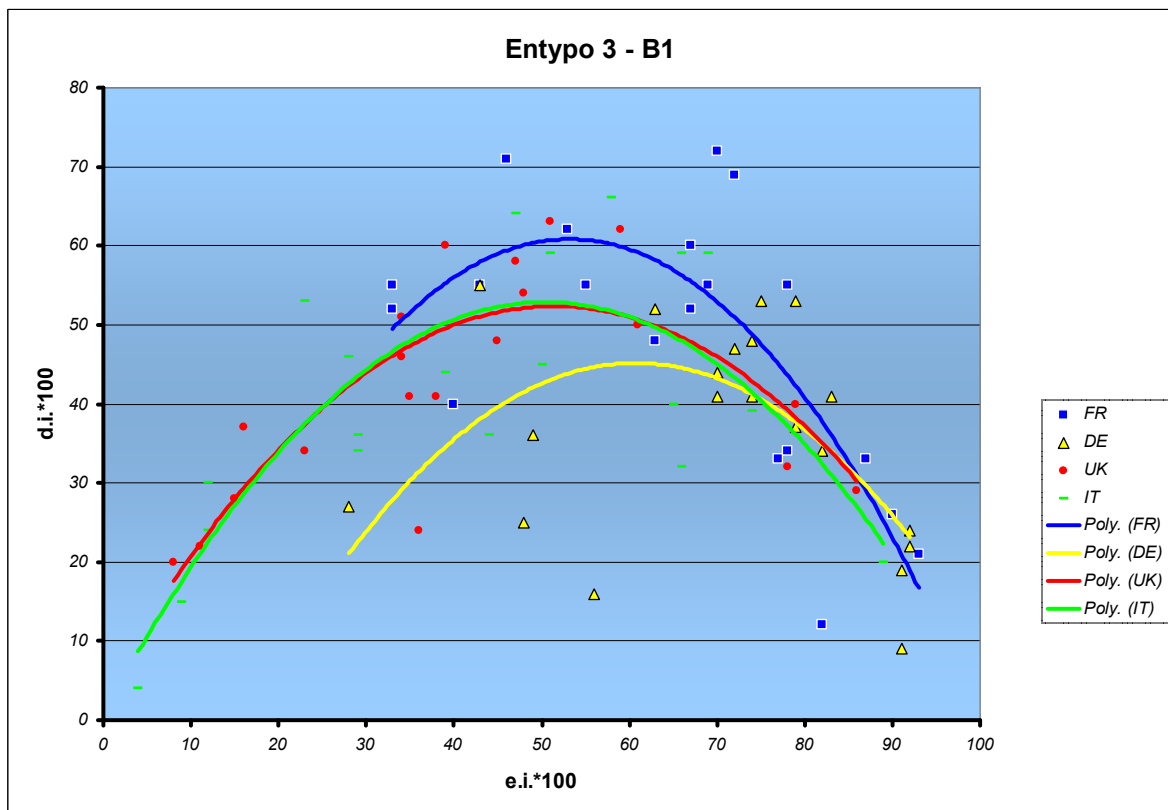
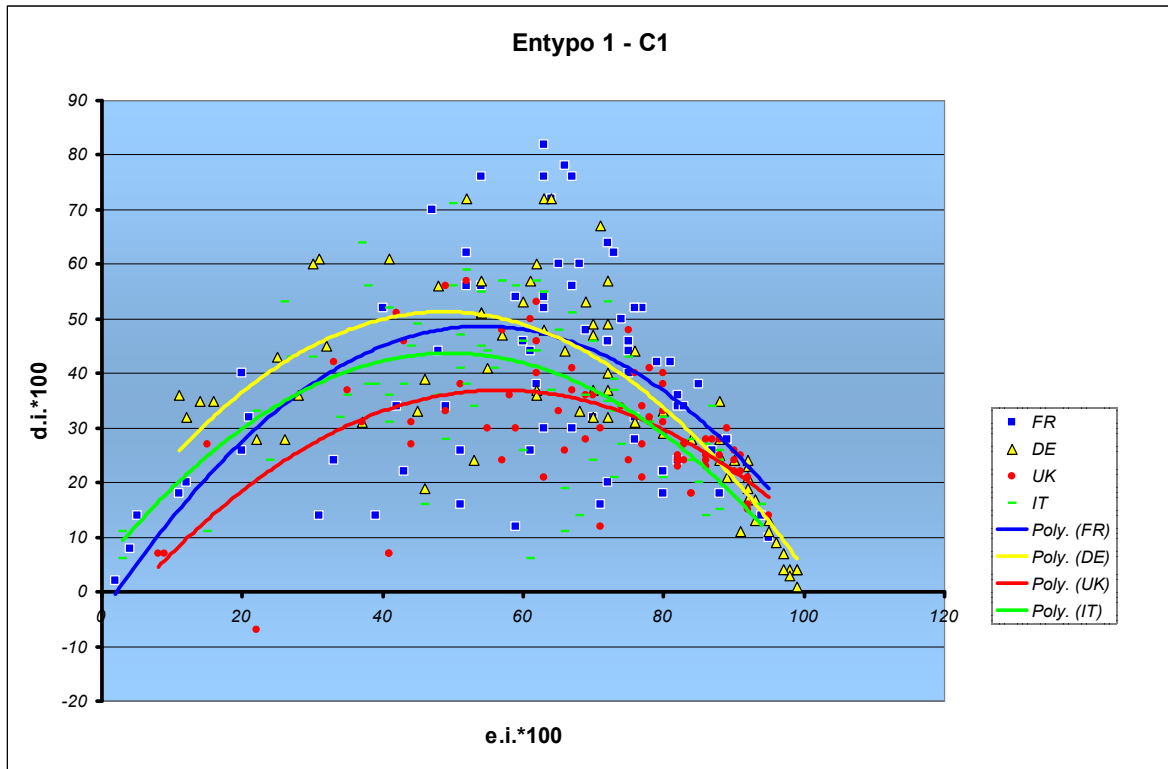
Ενότητα-Περίοδος-Επίπεδο		Αγγλική	Γαλλική	Γερμανική	Ιταλική
Πρώτη ενότητα	2007A Β1	41,20	30,98	40,18	39,36
	2007A Β2	40,07	37,05	30,95	32,72
	2007A Γ1	29,43	39,12	34,36	36,12
	2007B Β1	42,54	28,72	38,10	38,28
	2007B Β2	37,56	30,04	38,96	35,00
	2007B Γ1	36,10	34,70	39,98	33,80
	2008A Α1+2	40,34	36,24	33,54	26,64
	2008A Β1	43,98	36,00	45,12	39,72
	2008A Β2	40,80	31,32	39,46	33,86
	2008A Γ1	34,43	29,53	43,77	38,55
Τρίτη ενότητα	2007A Β1	42,00	48,00	36,20	40,25
	2007A Β2	34,72	37,56	33,24	32,52
	2007A Γ1	36,76	24,48	34,56	31,60
	2007B Β1	49,45	44,55	37,20	36,95
	2007B Β2	37,85	40,85	37,95	30,20
	2007B Γ1	31,52	41,08	37,40	29,20
	2008A Α1+2	46,10	48,40	46,15	42,60
	2008A Β1	47,45	38,95	41,20	39,20
	2008A Β2	36,95	36,60	36,55	43,30
	2008A Γ1	36,96	30,56	40,16	37,56
ΜΕΣΗ ΤΙΜΗ:		39,31	36,24	38,25	35,87
ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ:		4,99	6,26	3,87	4,39

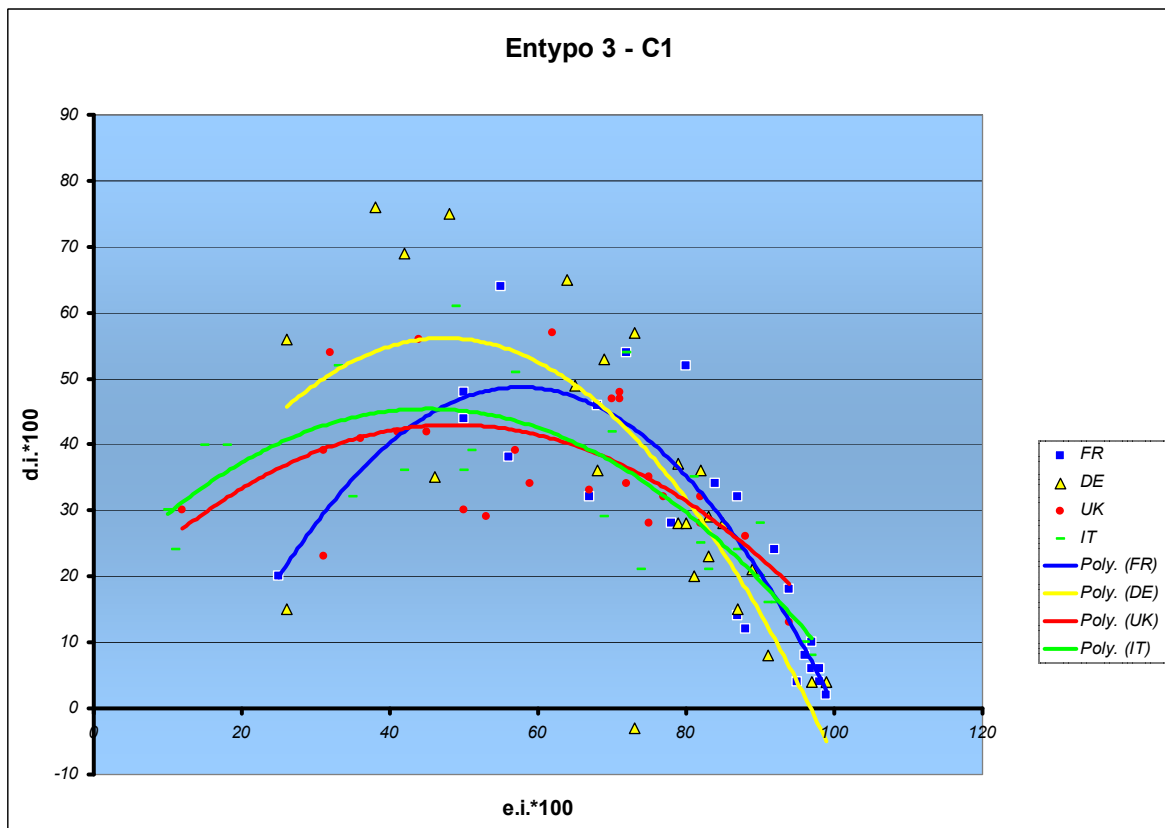
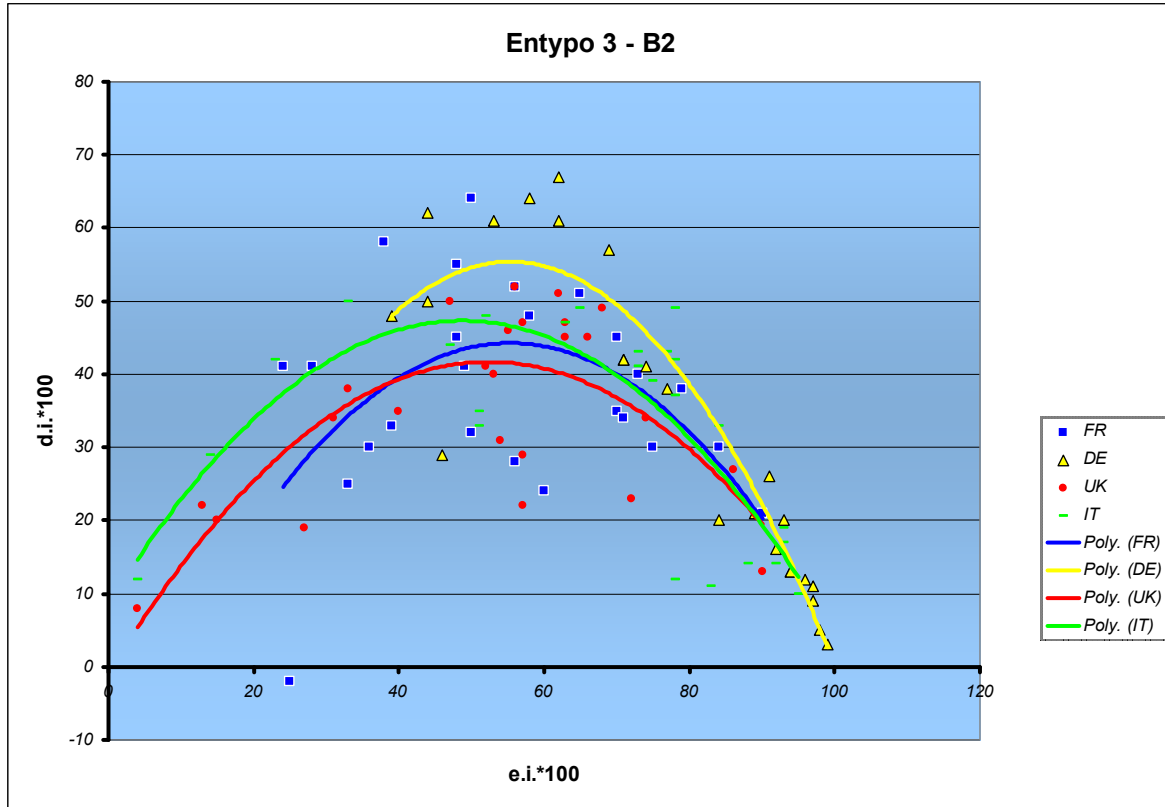
Το τελικό συμπέρασμα που μπορεί να εξαχθεί -στηριζόμενοι και σε στοιχεία που δεν παρουσιάζονται ούτε αναλύονται στο παρόν Παραδοτέο (για παράδειγμα σχετικά με την εγκυρότητα περιεχομένου, τη διαβαθμολογική αξιοπιστία, κτλ.)- είναι πως το Κρατικό Πιστοποιητικό Γλωσσομάθειας είναι ένα έγκυρο σύστημα πιστοποίησης, εφάμιλλο των «καλών» ευρωπαϊκών συστημάτων.

ΠΑΡΑΡΤΗΜΑ: Στικτά διαγράμματα

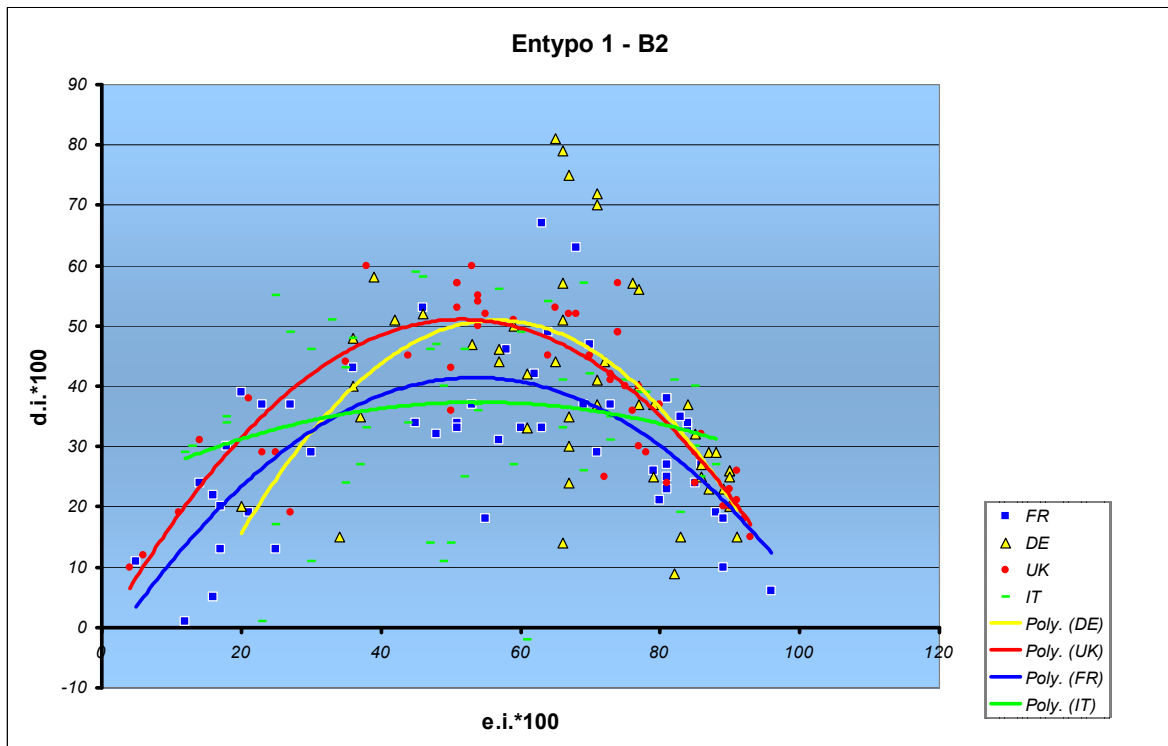
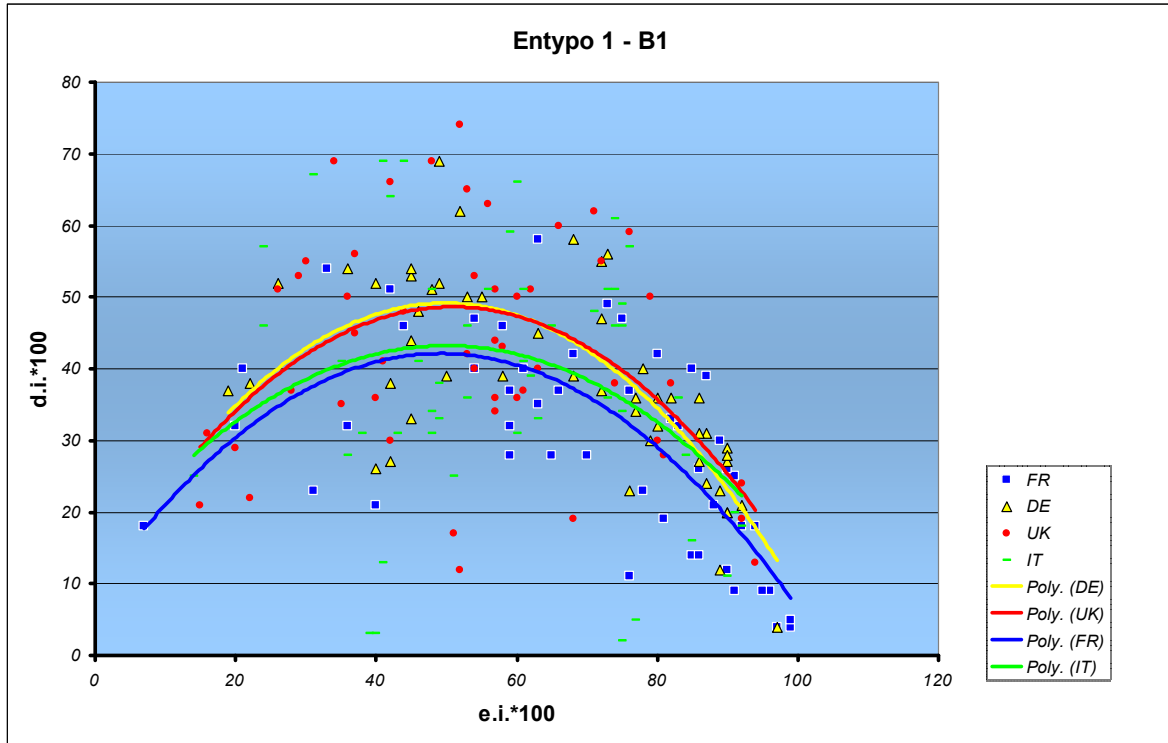
A. Μάιος 2007

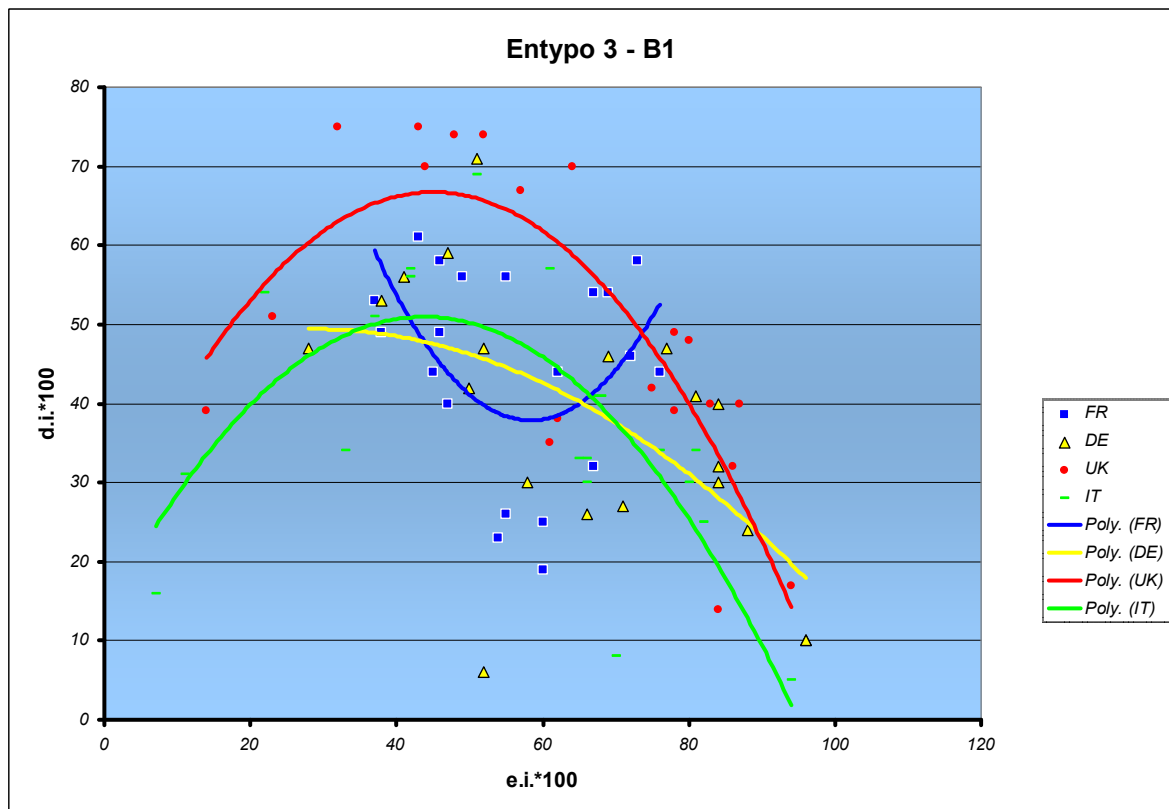
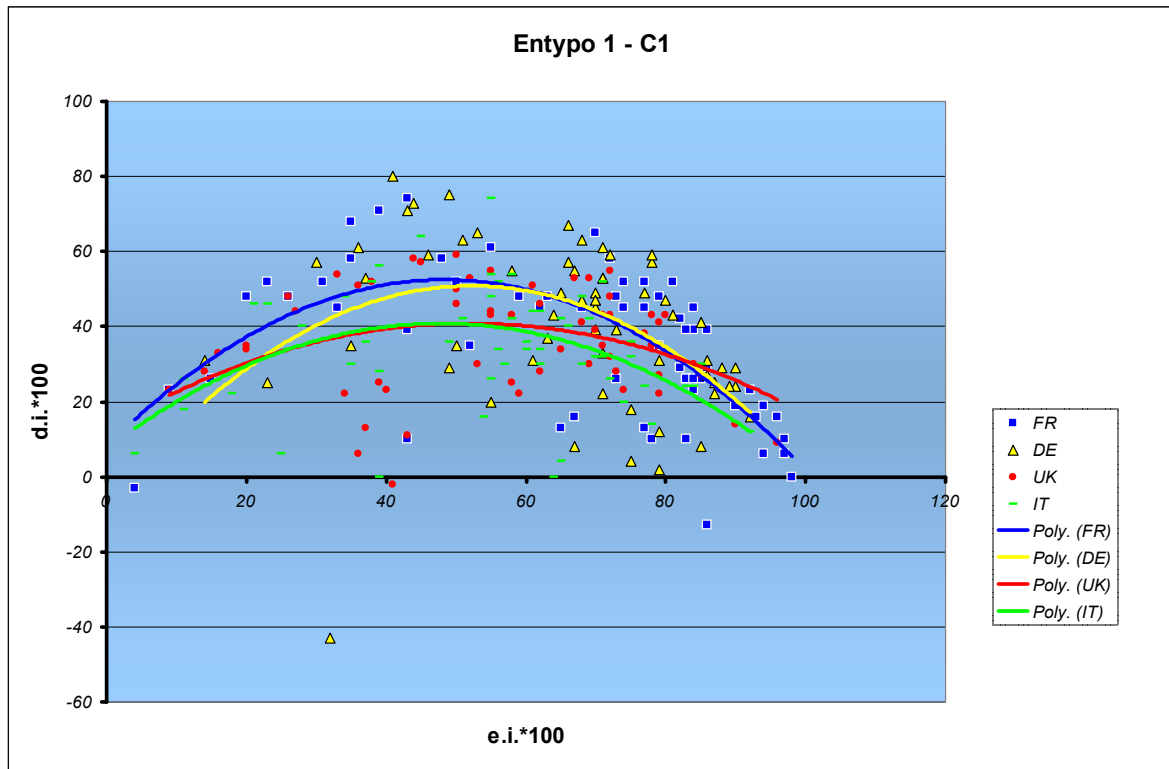


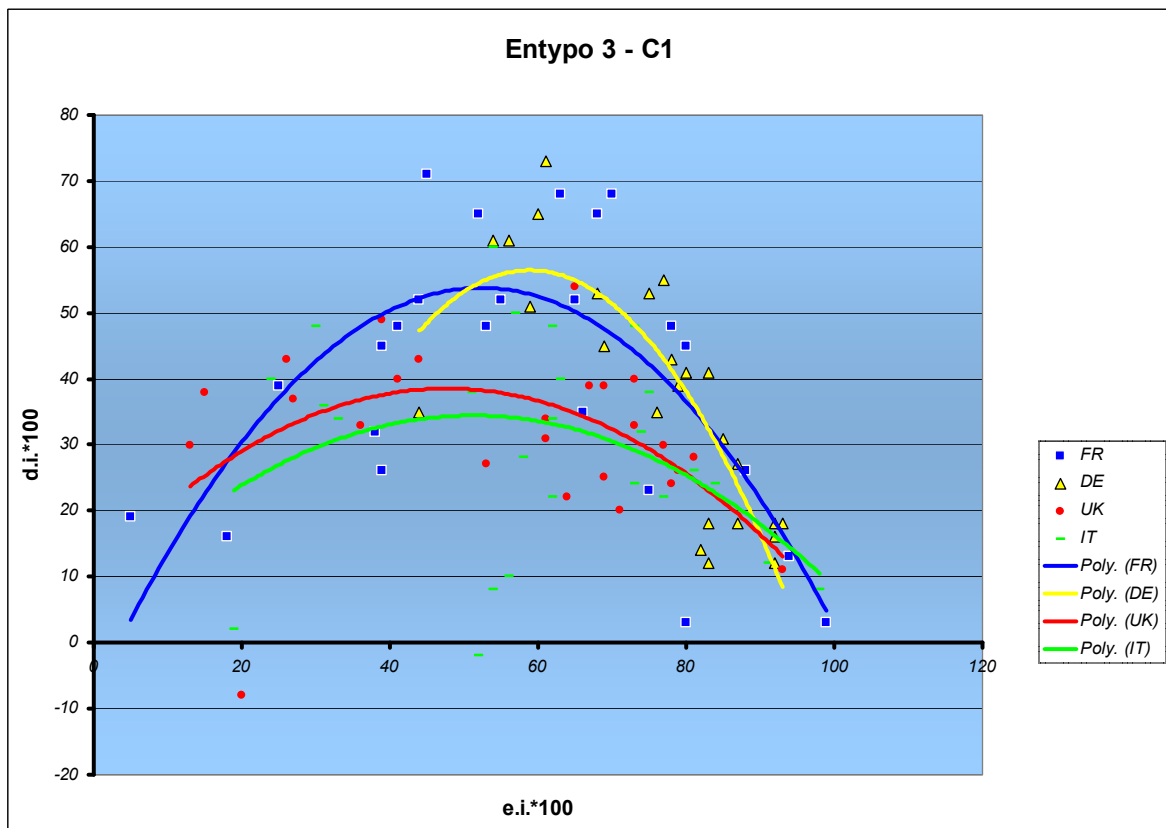
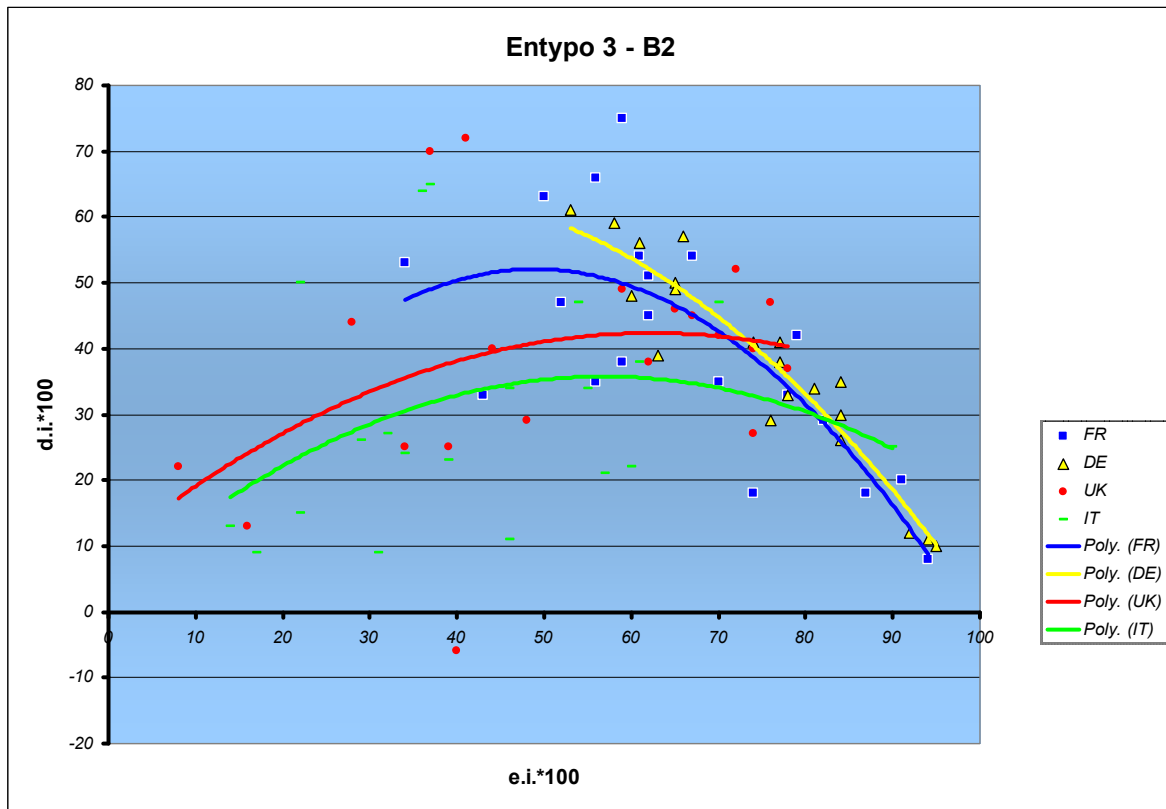




B. Νοέμβριος 2007







Γ. Μάιος 2008

