

**ΑΞΙΟΠΟΙΗΣΗ ΤΩΝ ΔΕΔΟΜΕΝΩΝ ΑΠΟ ΤΗΝ ΜΗΧΑΝΙΚΗ ΕΠΕΞΕΡΓΑΣΙΑ  
ΤΟΥ ΘΗΣΑΥΡΟΥ ΤΗΣ ΕΛΛΗΝΙΚΗΣ ΓΛΩΣΣΑΣ (TLG\*)**

ΓΕΩΡΓΙΟΣ ΜΠΑΜΠΙΝΙΩΤΗΣ

The author proposes a number of language processing programs based on the data base of Thesaurus Linguae Graecae. Proposed programs refer to textual and morphological distribution of data base material, to token verification, reverse thesaurus, earliest evidence thesaurus etc.

**1. Εισαγωγικά**

Έχοντας δουλέψει πρόσφατα σε μια σειρά προγραμμάτων αναλύσεως τής νέας ελληνικής γλώσσας μέσω Η/Υ, έχω εντυπωσιασθεί από τις δυνατότητες που μπορεί να προσφέρει η ηλεκτρονική επεξεργασία τής γλώσσας τών κειμένων.

Η επεξεργασία κειμένων έχει, ως γνωστόν, σήμερα προχωρήσει πολύ πιο πέρα από την απλή ηλεκτρονική καταγραφή τών δεδομένων σε ελάχιστο χώρο και την ανάκλησή τους σε ελάχιστο χρόνο (μνήμη), πέρα από στατιστικές πληροφορίες και ποικίλες ομαδοποιήσεις τών δεδομένων· έχει προχωρήσει σε πολύ σύνθετες και λεπτές μορφές επεξεργασίας, αυτές που προϋποθέτει λ.χ. η μηχανική μετάφραση (machine translation) ή η γραμματικο-συντακτική αναγνώριση (parsing) και παραγωγή τών συστατικών μιας γλώσσας.

Οι σκέψεις που θα εκθέσω εν συνεχεία αφορούν κυρίως σε ειδικότερα («μεσοπρόθεσμα» τρόπον τινά) προγράμματα επεξεργασίας τού κειμένου. Προϋποθέτουν ή και προχωρούν παράλληλα με την πρώτη φάση τής καταγραφής τών δεδομένων (την «τράπεζα δεδομένων»), αλλά δεν φθάνουν σε «μακροπρόθεσμα» —ας τα πούμε— προγράμματα σύνθετης επεξεργασίας τών δεδομένων, όπως θα ήταν λ.χ. η σύνταξη ενός

---

\* Το παρόν κείμενο είναι η (χωρίς προσθήκες ή συμπληρώσεις) ανακοίνωση τού γράφοντος στο συνέδριο τού «Θησαυρού τής Ελληνικής Γλώσσας» που πραγματοποιήθηκε στην Αθήνα (18-20 Ιουλίου 1987).

ερμηνευτικού λεξικού ή ενός λεξικού συνωνύμων τής αρχαίας ή τής νέας ή ολόκληρης τής Ελληνικής.

## 2. Τράπεζα λεξικών τύπων: Η προϋπόθεση

Όπως ελέχθη ήδη, βασική προϋπόθεση κάθε άλλου σύνθετου ή συνθετότερου προγράμματος είναι ο καταρτισμός μιας «τράπεζας λεξικών δεδομένων», η καταγραφή στον Η/Υ τών «λέξεων» (ορθότερα τών «λεξικών τύπων» ή, ακόμη ορθότερα, τών «λεξικών δειγμάτων» [tokens]) που απαντούν στα κείμενα τα οποία μας ενδιαφέρουν. Η τράπεζα τών λεξικών τύπων, στοιχείων δηλ. μηχανικώς αναγνώσιμων (machine readable) και επεξεργάσιμων (machine processed), αποτελεί την πρώτη και βασική ύλη οποιασδήποτε περαιτέρω επεξεργασίας. Το λεξικό (ή «το πρόγραμμα τού λεξικού»), όπως συνήθως λέγεται, είναι η προϋπόθεση τών λοιπών προγραμμάτων, η δε συγκρότησή του έχει τα δικά της προβλήματα (κείμενα που θα μηχανογραφηθούν, πληροφορίες που θα τα συνοδεύουν κ.λπ.), αυτά που αντιμετωπίστηκαν στην συγκρότηση τού Θησαυρού τής αρχαίας ελληνικής γλώσσας και που μέλλει να αντιμετωπιστούν και στον θησαυρό τής βυζαντινής και νεότερης Ελληνικής.

## 3. Προγραμματισμός παράλληλων προγραμμάτων

Η τράπεζα λεξικών τύπων τής αρχαίας γλώσσας έχει καταρτισθεί. Είναι το «πρόγραμμα TLG» τού Irvine. Μένει να συνταχθεί —κι αυτό είναι το θέμα αυτού τού Συνεδρίου— η αντίστοιχη τράπεζα τής μεσαιωνικής (βυζαντινής) και τής νέας ελληνικής γλώσσας. Ακριβώς σ' αυτή την προκαταρκτική φάση τού έργου μπορούν να προβλεφθούν και να προετοιμασθούν ορισμένα συναφή επιμέρους προγράμματα, στα οποία θα αναφερθώ εν συνεχεία. Για την εφαρμογή αυτών τών προγραμμάτων απαιτείται κατά την καταγραφή τών λεξικών τύπων να δίνονται συγχρόνως ορισμένες πληροφορίες (στοιχεία) που θα αξιοποιηθούν κατόπιν στα επί μέρους (ειδικά) προγράμματα.

Αν δεν μελετηθεί και προετοιμασθεί εξ αρχής το θέμα τής πολύπλευρης αξιοποίησης τών λεξικών δεδομένων σε επί μέρους προγράμματα, τότε κάθε ένα από αυτά θα απαιτεί σχεδόν νέα, εξ αρχής εργασία, που θα καθιστά το πρόγραμμα —εκ τών υστέρων— αντιοικονομικό. Επομένως —κι αυτό ισχύει για πολλά προγράμματα— θα πρέπει αυτά τα προγράμματα να αποφασισθούν από την αρχή και να προετοιμασθούν καταλλήλως.

Οι σκέψεις που θα ακολουθήσουν αφορούν σε μερικά τέτοια προγράμματα.

## 4. Γενικό πρόγραμμα

Το κεντρικό («βασικό» ή «γενικό») πρόγραμμα τόσο τής αρχαίας που είναι έτοιμο όσο και τών νεότερων φάσεων τής ελληνικής γραμματείας και γλώσσας που προγραμματίζονται αποτελεί η αποθησαύριση στον Η/Υ τού συνόλου τών μαρτυρουμένων στα γραπτά κείμενα λεξικών τύπων. (Ας τονισθεί, επ' ευκαιρία, πως ο μηχανικής επεξεργασίας Θησαυρός αποτελεί τράπεζα λεξικών τύπων, δηλ. μαρτυριών —όχι ερμηνευμάτων (σημασιών κ.λπ.), που θα απαιτούσε άλλο, πολύ πιο σύνθετο πρόγραμμα).

Συγκεκριμένα το γενικό πρόγραμμα ενός θησαυρού καθ' εαυτό περιλαμβάνει, κανονικώς, τα εξής:

(i) Το σύνολο των λεξικών τύπων τής αποθησαυριζόμενης γλώσσας (π.χ. *γράφω* Θουκ. ..., *έγραψε* Ηρόδ. ..., *γεγραμμένος*...).

(ii) Το σύνολο των μαρτυριών κάθε λεξικού τύπου (των χωρίων δηλ. των κειμένων όπου απαντά).

(iii) Το σύνολο των περιβαλλόντων (κυμαινόμενης εκτάσεως) κάθε λεξικού τύπου.

(iv) Την συχνότητα εμφάνισης κάθε λεξικού τύπου.

Οίκοθεν νοείται ότι τα 4 αυτά είδη πληροφοριών είναι κι αυτά που θα υπάρχουν στο γενικό πρόγραμμα τής βυζαντινής και τής νεοελληνικής γραμματείας/γλώσσας. Περαιτέρω προβλέψεις οδηγούν στην σύνταξη ειδικών προγραμμάτων που μπορούν να απορρέουν από έναν γλωσσικό θησαυρό χωρίς και να υπονοούνται αναγκαστικά απ' αυτόν.

### 5. Προτεινόμενα ειδικά προγράμματα

Χωρίς να επιχειρούμε εδώ εξαντλητική εξέταση όλων των δυνατών προγραμμάτων που μπορούν να προκύψουν από το γενικό πρόγραμμα ενός Θησαυρού τής γλώσσας, θα προτείνουμε ορισμένα ειδικά προγράμματα άμεσης προτεραιότητας, αφήνοντας για άλλους ή για άλλη ευκαιρία άλλα δυνατά προγράμματα.

Συγκεκριμένα προτείνουμε τα εξής 12 ειδικά προγράμματα, τα οποία για τα βυζαντινά και τα νέα ελληνικά μπορούν να προβλεφθούν εξ αρχής και να «τρέχουν» παράλληλα προς το γενικό πρόγραμμα (για το συντελεσμένο πια πρόγραμμα τής αρχαίας γλώσσας θα πρέπει να διατρέξει κανείς εκ νέου όλο το αποθησαυρισμένο υλικό και να ξεκινήσει στην πραγματικότητα εξ αρχής).

(i) *Πρόγραμμα κειμενικής και μορφολογικής κατανομής των λεξημάτων κατά λήμματα* (Textual and morphological distribution program). Πρόγραμμα εμφάνισης τής συνολικής εικόνας ενός λήμματος, τού συνόλου δηλ. των μαρτυρουμένων λεξικών τύπων (ή, αλλιώς, τού «μορφολογικού παραδείγματος») ενός λήματος-μάννας. Αντί δηλ. οι πληροφορίες που χρειάζεται ο ερευνητής να παρέχονται μόνο διά τής αναζητήσεως ενός συγκεκριμένου λεξικού τύπου (τού τύπου *έποιησαν* λ.χ.), με ένα τέτοιο πρόγραμμα «λημματισμού» θα μπορεί ο χρήστης να ανάγεται απευθείας και στο λήμμα (λ.χ. το λ. *ποιῶ*) και να έχει συγχρόνως δύο είδη πληροφοριών:

α) Την συνολική στατιστική και κειμενική εμφάνιση τού λ. *ποιῶ*.

β) Αναλυτικώς το σύνολο των μαρτυρουμένων λεξικών τύπων ενός λήματος (την κατανομή του σε τύπους) μαζί με τα χωρία όπου απαντούν και τις στατιστικές ενδείξεις τους.

Με το πρόγραμμα αυτό μπορεί ο χρήστης (μελετητής, ερευνητής, δάσκαλος κ.λπ.) να αποκτά συνολική εικόνα τής κειμενικής και μορφολογικής κατανομής ενός λεξήματος μαζί με τις ανάλογες (συνολική και μερική) στατιστικές ενδείξεις.

(ii) *Πρόγραμμα αναγνώρισεως ενός λεξικού τύπου εν κειμένω* (token verification program). Αναγνώριση ή μη αναγνώριση (με σχετική σήμανση) οποιουδήποτε γλωσσικού/λεξικού τύπου (οποιασδήποτε λέξεως) εν κειμένω. Κατά την εμφάνιση στην οθόνη τού τερματικού ενός κειμένου θα ελέγχεται μέσω τού προγράμματος αυτού αν μια λέξη ως λεξικός τύπος (π.χ. *έποιησαντο*, *γράφουν*, *εμένα* κ.λπ.) απαντά ή δεν απαντά στα κείμενα τής συγκεκριμένης περιόδου (αρχαίας, βυζαντινής ή νέας) για την οποία έχει γίνει το πρόγραμμα.

Με το πρόγραμμα αυτό επιτυγχάνεται αυτόματος έλεγχος τής κειμενικής ή μη εμφάνισης μιας λέξεως σε ορισμένη γλωσσική περίοδο.

(iii) *Πρόγραμμα θησαυρού πληροφοριών κριτικού υπομνήματος* (Apparatus criticus program). Το πρόγραμμα αυτό συγκεντρώνει και παρέχει πληροφορίες για γραφές, διορθώσεις, αναγνώσεις των διαφόρων λέξεων α) κατά λήμμα, β) κατά συγγραφέα, κείμενο και χωρίο τού κειμένου.

Η χρησιμότητα ενός τέτοιου προγράμματος, που θα μπορεί να συγκεντρώνει και να παρέχει κατά συγγραφέα και κείμενο τις υπάρχουσες γραφές, είναι προφανής.

(iv) *Πρόγραμμα αντίστροφου θησαυρού τής ελληνικής γλώσσας* (Reverse thesaurus program). Πρόγραμμα αντίστροφης κατατάξεως τού θησαυρού τής ελληνικής γλώσσας, κατά λήμματα και τύπους, συνολικά αλλά και μερικά (κατά περίοδο ή συγγραφέα). Σύμφωνα με το πρόγραμμα αυτό θα μπορεί ο χρήστης να βρίσκει έτοιμο και συγκεντρωμένο το υλικό για τη συμπλήρωση επιγραφικών ή άλλων ελλιπώς μαρτυρουμένων κειμένων τών οποίων σώζεται μόνο το τέλος τών λέξεων. Ακόμη ένα τέτοιο πρόγραμμα παρέχει ανάγλυφη εικόνα τής μορφολογικής δομής και κατανομής μιας γλώσσας, η οποία συμπληρούμενη με στατιστικές πληροφορίες χρήσεως μπορεί να αποβεί πολύτιμο εργαλείο στη διδασκαλία τής γλώσσας, στην επιλογή υλικού για τη σύνταξη βιβλίων κ.λπ.

Επ' ευκαιρία, πολύτιμο —χωρίς να αποτελεί και ιδιαίτερο ή δύσκολο πρόγραμμα— αφού οι στατιστικές πληροφορίες είναι απλό θέμα για τον Η/Υ, είναι η αντίστροφη κατάταξη τών λεξικών τύπων τού Θησαυρού σε πίνακα κατά στατιστική συχνότητα κάθε λεξικού τύπου, γεγονός που είναι και πάλι πρωταρχικό στην εκπαιδευτική χρήση τών κειμένων.

(v) *Πρόγραμμα ειδικών λεξιλογίων κατά συγγραφέα και κατά κείμενο* (Author and text vocabulary program). Συγκρότηση ειδικών λεξικών (λεξιλογίων) κατά συγγραφέα και κείμενο με βάση τα λήμματα και την μορφολογική κατανομή τους.

Μ' αυτό το πρόγραμμα θα έχουμε κάθε λήμμα (ή λεξικό τύπο) σε πληρέστερη μορφή με αναφορά όλων τών χωρίων όπου απαντά και στατιστικές ενδείξεις. Τέτοια λεξικά υπάρχουν ήδη για αρκετούς αρχαίους συγγραφείς, αλλά δεν είναι πάντα στον τύπο τού Θησαυρού (με πλήρη συγκεντρωτική αναφορά όλων τών μαρτυριών). Ακόμη τέτοια λεξικά μπορούν να εφοδιασθούν και να συσχετισθούν εύκολα με πληροφορίες για τα Σχόλια που υπάρχουν για κάθε λεξικό στοιχείο ή χωρίο κειμένου τού συγγραφέα.

(vi) *Πρόγραμμα ειδολογικών λεξικών* (Literary genre thesaurus program). Εύκολα με κατάλληλες ομαδοποιήσεις συγγραφέων και κειμένων μπορούν μ' αυτό το πρόγραμμα να παραχθούν *ειδολογικά λεξικά (θησαυροί)*: λ.χ. λεξικό (θησαυρός) λυρικών ποιητών, θησαυρός ιστορικών, θησαυρός τραγικών κ.ο.κ.

Η διαφορά πάλι από ορισμένα υπάρχοντα ήδη λεξικά είναι ακριβώς ότι θα πρόκειται περί ειδολογικών λεξιλογικών θησαυρών με το σύνολο τών πληροφοριών για την εμφάνιση μιας λέξεως και όχι περί ενδεικτικών, αντιπροσωπευτικών ή δειγματοληπτικών μαρτυριών όπως συχνά γίνεται.

(vii) *Πρόγραμμα χρονολογικών λεξικών* (Chronological thesaurus program). Πρόγραμμα συντάξεως θησαυρών κατά χρονικές περιόδους: π.χ. θησαυρός τής ελληνικής γλώσσας τού 6ου αι., τού 3ου αι., τού 2ου μ.Χ. αι. κ.ο.κ.

Με ένα τέτοιο πρόγραμμα θα είμαστε σε θέση να γνωρίζουμε τη *λεξιλογική σύνθεση τής γλώσσας* —άρα και την εξέλιξή της— κατά διάφορες χρονικές περιόδους.

(viii) *Πρόγραμμα θησαυρού κυρίων ονομάτων* (Proper names thesaurus program). Σύνταξη θησαυρού με όλα τα κύρια ονόματα κάθε περιόδου τής Ελληνικής: αρχαίας,

βυζαντινής και νέας —ή ολόκληρης τής Ελληνικής.

Στο ίδιο πρόγραμμα είναι δυνατόν να παραχθούν επιμέρους *θησαυροί τοπωνυμίων, ανθρωπωνυμίων, άλλων κυριωνυμίων* κ.λπ. Από ένα τέτοιο πρόγραμμα μπορεί αμέσως ο χρήστης να πληροφορηθεί ποιοί συγγραφείς ποιών περιόδων γνωρίζουν αυτό ή εκείνο το τοπωνύμιο, αυτό ή εκείνο το ανθρωπωνύμιο.

(ix) *Πρόγραμμα θησαυρού α' μνείας* (First/earliest evidence thesaurus program). Πρόγραμμα με πληροφορίες για την α' μνεία, πού και πότε πρωτοαπαντά κάθε λέξη ή λεξικός τύπος.

Από ένα τέτοιο πρόγραμμα θα είναι στους μελετητές προσιτές οι πληροφορίες για την ιστορική προέλευση κάθε λεξήματος.

(x) *Πρόγραμμα θησαυρού παράλληλων χωρίων* (Parallel passages thesaurus program). Πρόγραμμα που θα εμφανίζει (με ληματογραφική βάση την αρχή του χωρίου ή το κύριο λεξιλογικό στοιχείο του) ποιά είναι τα παράλληλα χωρία (και σε ποιά έκταση παράλληλα) που απαντούν σε διαφόρους συγγραφείς.

Πόσο διευκολύνεται από ένα τέτοιο πρόγραμμα η μελέτη των επιδράσεων που αποτελούν κύριο αντικείμενο των φιλολογικών και γλωσσικών σπουδών δεν χρειάζεται να σχολιασθεί. Το γενικότερο και μόνιμο πρόβλημα των γλωσσικών-υφολογικών επιδράσεων μπορεί με αρκετή ασφάλεια να μελετηθεί μέσα από το πρόγραμμα αυτό.

(xi) *Πρόγραμμα θησαυρού συντακτικών (ή συνταγματικών) σχέσεων* (Syntactic/syntagmatic relations thesaurus program). Σχέσεις όπως η *σύνταξη των ρημάτων* (με αιτιατική ή γενική ή δοτική ή με δύο πτώσεις κ.λπ.) ή η *σύνταξη των ονομάτων*, η *σύνταξη των προθέσεων* κ.τ.ό. μπορούν εύκολα να προκύψουν από αυτό το πρόγραμμα.

Η χρησιμότητα ενός τέτοιου προγράμματος για ερευνητικούς αλλά και για εκπαιδευτικούς σκοπούς είναι εμφανής.

(xii) *Πρόγραμμα θησαυρού λεξιλογικών κατηγοριών* (Parts of speech thesaurus program). Μέσα από αυτό το πρόγραμμα θα προκύψουν οι κατηγοριοποιήσεις του θησαυρού *κατά μέρη του λόγου: θησαυρός ουσιαστικών, θησαυρός επιθέτων, θησαυρός ρημάτων* κ.ο.κ., όπου για κάθε στοιχείο (ουσιαστικό, επίθετο, ρήμα κ.λπ.) μπορούν με κατάλληλο προγραμματισμό να δίδονται ποικίλες πληροφορίες (μετβ. — αμετβ. ρήμα κ.λπ.).

## 6. Προγράμματα σύνθετης υφής

Τα προγράμματα που προτείναμε ως τώρα ανταποκρίνονται σε άμεσες χρήσεις του αποθησαυριζομένου λεξικού υλικού· αποτελούν, τρόπον τινά, άμεσα ή παράλληλα προς το γενικό προγράμματα. Ωστόσο, όπως αναμένεται, από τα άμεσα αυτά προγράμματα και με αφετηρία πάντα το γενικό πρόγραμμα είναι δυνατόν να προέλθουν επεξεργασμένα προγράμματα σύνθετης υφής, προγράμματα που άμεσα ή έμμεσα προϋποθέτουν τα μνημονευθέντα.

Αναφέρω ενδεικτικά μερικά από αυτά τα προγράμματα:

(i) *Ερμηνευτικά λεξικά*. Είναι αυτονόητο ότι ένα ερμηνευτικό λεξικό οποιασδήποτε περιόδου (αρχαίο, βυζαντινό ή νεότερο), για να είναι επαρκές και αξιόπιστο, προϋποθέτει την γνώση και αξιοποίηση του υλικού του αντίστοιχου θησαυρού δεδομένων που θα επιτρέψει στον συντάκτη του ερμηνευτικού λεξικού να βρει, να ταξινομήσει, να επαληθεύσει/διαψεύσει/αναθεωρήσει τα παραδοσιακά ερμηνεύματα και, κυρίως, να τα συμπληρώσει από τις σημασίες-χρήσεις που θα προκύψουν από την

πληρέστερη και συνολική θεώρηση τού γλωσσικού υλικού. Η λεξικογραφική εργασία που γινόταν μέχρι σήμερα —συχνά καρπός μόχθου ενός και μόνον ανθρώπου— ήταν φυσικό να έχει επιλεκτικό, αποσπασματικό και περιορισμένο σε έκταση υλικού χαρακτήρα. Το γεγονός ότι με τον ηλεκτρονικό θησαυρό ο μελετητής (εν προκειμένω ο λεξικογράφος) μπορεί ανά πάσαν στιγμή να έχει στη διάθεσή του ό,τι πληροφορία τον ενδιαφέρει (να ανακαλεί και να ελέγχει το χωρίο που τον ενδιαφέρει μέσα στην οθόνη τού τερματικού του) και μάλιστα σε όλη την έκταση τού υλικού τής περιόδου που μελετά, κι ακόμη να μπορεί —εφόσον προηγηθεί η ηλεκτρονική καταγραφή του— να ανακαλεί και να συγκρίνει το υλικό του με στοιχεία άλλων περιόδων, το γεγονός αυτό αλλάζει άρδην την λεξικογραφική εργασία στις μέρες μας.

Το ερμηνευτικό λεξικό, που αποτελεί μια από τις πιο σύνθετες μορφές λεξικογραφικών έργων, αφού η σύνταξή του απαιτεί, στην πραγματικότητα, να ληφθούν υπόψιν όλες οι *παραδειγματικές* σχέσεις (συνωνύμων, αντωνύμων, υπωνύμων) καθώς και οι *συνταγματικές* σχέσεις (σχέσεις συντακτικού συμπληρώματος τού ρήματος, προσδιορισμού τών ουσιαστικών και τών ρημάτων, επιρρηματικών συνόλων κ.λπ. κ.λπ.), είναι δυνατόν να συνταχθεί επαρκώς μόνον εφόσον είναι στον λεξικογράφο διαθέσιμο το υλικό ενός ηλεκτρονικού θησαυρού της γλώσσας.

(ii) *Ειδικά λεξικά*: *Λεξικό επιθέτων* (επιθετικών διορισμών που χρησιμοποιούνται στατιστικώς συχνότερα για κάθε ουσιαστικό): *Λεξικό φράσεων* (παγίων φράσεων που χρησιμοποιούνται σε μια περίοδο τής γλώσσας: τύπος: *εκ των ενόντων*): *Υφολογικό λεξικό* (λεξικό αποκλειστικώς τών χρήσεων, τών παγίων δυνατών συνταγματικών σχέσεων κάθε λεξήματος, που χρησιμοποιούνται σε ορισμένη χρονική περίοδο τής γλώσσας ή σε ορισμένο συγγραφέα ή σε κύκλο συγγραφέων κ.ο.κ.).

Ενδεικτικά, όπως είπαμε, αναφέρθηκαν εδώ μερικές από τις «δευτερογενείς» εφαρμογές ενός θησαυρού τής γλώσσας, αυτές που μπορεί να ακολουθήσουν το γενικό βασικό πρόγραμμα και άλλα δυνατά παράλληλα προγράμματα. Οι ειδικοί μελετητές, προπάντων οι φιλόλογοι αλλά και όλοι όσοι ασχολούνται με τη γλώσσα τών κειμένων, μπορούν να σκεφθούν άλλες χρήσεις/εφαρμογές τού Θησαυρού που μέσα από ειδικά (απλά ή σύνθετα) προγράμματα θα αξιοποιούν το πολύτιμο υλικό ενός ηλεκτρονικού θησαυρού με τεράστιες δυνατότητες αποθηκείσεως πληροφοριών (μνήμη) και ελάχιστους χρόνους ανακλήσεώς τους.

## 7. Επιλεγόμενα

Οι Η/Υ —δεν το έχουμε ίσως ακόμη συνειδητοποιήσει οι επιστήμονες τών ανθρωπιστικών σπουδών— αποτελούν το τελειότερο εργαλείο, το επαναλαμβάνω *εργαλείο*, που επινόησε η σκέψη τού ανθρώπου όχι μόνο για τεχνολογικές εφαρμογές (όσο κι αν οι εφαρμογές τών Η/Υ στον χώρο αυτόν είναι πράγματι εντυπωσιακές), αλλά για να υπηρετήσει κάθε υπολογιστική-ταξινομική-συνδυαστική ανάγκη τού ανθρώπου σε κάθε τομέα τής δραστηριότητάς του.

Σήμερα η ηλεκτρονική (με Η/Υ) επεξεργασία τών κειμένων αποτελεί πραγματική επανάσταση στις δυνατότητες επικοινωνίας τού ανθρώπου. Υπάρχουν ήδη προγράμματα για την αγγλική γλώσσα που διορθώνουν αυτόματα τα ορθογραφικά λάθη ενός κειμένου, που επισημαίνουν τα γραμματικά και τα συντακτικά ακόμη λάθη, ενώ συγχρόνως τον βοηθούν να τα διορθώσει με υποδείξεις που γίνονται στην οθόνη τού τερματικού τού χρήστη. Έχει επιτευχθεί —το είπα ήδη— «αυτόματη μετάφραση κειμένων» (με ατέλειες ακόμη), ενώ ήδη (σε προγράμματα τής Αγγλικής) παρέχεται

στον χρήστη πολύτιμη λεξιλογική βοήθεια μέσω προγραμμάτων συνωνύμων, υφολογικών χρήσεων, ομοήχων κ.λπ.

Η ηλεκτρονική επεξεργασία τών κειμένων, τέλος, έχει περάσει και στον εκπαιδευτικό χώρο με ποικίλες εφαρμογές σε σχολικά βιβλία και ασκήσεις στην τάξη όσον αφορά στην σύνταξη και επεξεργασία κειμένων.

Το μεγάλο πλεονέκτημα τού Η/Υ είναι η οικονομία δυνάμεων στην συγκέντρωση, την αποθήκευση την ταξινόμηση και την ανάκληση τών πληροφοριών. Θυμηθείτε, ας θυμηθούμε όλοι μας, πόσο χρόνο τής ζωής μας έχουμε δαπανήσει ψάχνοντας μια λέξη, έναν τύπο, ένα όνομα, μια φράση, μια σύνταξη σε χιλιάδες σελίδων, σε ατέλειωτους τόμους κειμένων... Κάνοντας οικονομία μνήμης, έχουμε χρόνο και δυνατότητες για καλύτερη αξιοποίηση τής κρίσης μας.

Ιδίως οι φιλόλογοι, *homines suspiciosi* εκ φύσεως και εξ επαγγέλματος, μπορούν να επινοήσουν χρήσεις και εφαρμογές τού Η/Υ που θα λύσουν με ακρίβεια, με ταχύτητα και, προπάντων, με αντικειμενική τεκμηρίωση και αξιοπιστία ποικίλα φιλολογικά ζητήματα τών κειμένων.

Τέλος, η μεγάλη αδυναμία μας στην Ελλάδα, ο καημός και το μεράκι μας, η βαριά κληρονομιά, το μεγάλο πρόβλημα αλλά και το μεγάλο όπλο μας, η γλώσσα μας, η ελληνική γλώσσα μπορεί —μέσα από τους ηλεκτρονικούς θησαυρούς τών επί μέρους φάσεων της Ελληνικής— να μελετηθεί σωστά και να αξιοποιηθεί (επικοινωνιακά, εκπαιδευτικά, μορφωτικά, εθνικά) στον μεγαλύτερο δυνατό βαθμό. Ό,τι ονομάζουμε *ενιαία* ελληνική γλώσσα, νοώντας μ' αυτό την διαχρονική ενότητα στους δομικούς αρμούς, στη γραπτή παράδοση και στην αδιάκοπη επί 20 αιώνες προφορική παράδοση τής γλώσσας μας, μπορεί άριστα να ερευνηθεί και να αξιοποιηθεί με την ηλεκτρονική επεξεργασία τών γλωσσικών κειμένων μας που συχνά τα χωρίζουν τόσοι αιώνες, και μαζί με τις τεράστιες τεχνικές δυνατότητες ανευρέσεως, συνδέσεως και, κυρίως, *αναλύσεως* που προσφέρει σήμερα η ηλεκτρονική σπουδή τών κειμένων μέσα από τον θησαυρό τής γλώσσας.

Γ. Μπαμπινιώτης  
Πανεπιστήμιο Αθηνών  
Τομέας Γλωσσολογίας