ΕΛΕΤΟ – 2ο Συνέδριο «Ελληνική Γλώσσα και Ορολογία»
Αθήνα, 21-23 Οκτωβρίου 1999

ΑΥΤΟΜΑΤΗ ΔΗΜΙΟΥΡΓΙΑ ΜΟΡΦΟΛΟΓΙΚΩΝ ΛΕΞΙΚΩΝ ΜΕ ΟΡΟΛΟΓΙΚΟ ΠΛΟΥΤΟ ΒΑΣΕΙ ΚΕΙΜΕΝΩΝ ΕΝΤΑΣΕΩΣ ΟΡΩΝ

Γ. Ταμπουρατζής, Γ. Καραγιάννης

ΠΕΡΙΛΗΨΗ

Η συμβατική μέθοδος κατασκευής λεξικών απαιτεί την καταγραφή όλων των μορφών των λέξεων της αντίστοιχης γλώσσας. Η ολοκλήρωση της μεθόδου αυτής απαιτεί εργασία που μετριέται σε ανθρωποέτη, ενώ είναι πρακτικά αδύνατο να καλυφθεί το σύνολο των λέξεων της γλώσσας. Ειδικότερα στην περίπτωση κειμένων εντάσεως όρων, όπου οι χρησιμοποιούμενες λέξεις εστιάζονται σε μία ορισμένη θεματική περιοχή, τα υπάρχοντα μορφολογικά λεξικά γενικής χρήσης δεν παρέχουν επαρκή κάλυψη. Στην ανακοίνωση αυτή προτείνεται μία μεθοδολογία, βασισμένη στην τεχνική της ταύτισης-και-απόκρυψης σε συνδυασμό με γραμματικούς περιορισμούς, για την παραγωγή εξειδικευμένων μορφολογικών λεξικών από κείμενα εντάσεως όρων. Πειραματικές δοκιμές δείχνουν ότι η μέθοδος αυτή δίνει αποτελέσματα που χαρακτηρίζονται από υψηλή ακρίβεια.

AUTOMATED CONSTRUCTION OF MORPHOLOGICAL LEXICA POSESSING TERMINOLOGY WEALTH ON THE BASIS OF TERM-INTENSIVE DOCUMENTS

ABSTRACT

The established method for generating morphological lexica requires the creation of a list of possible wordforms in the respective language. This needs several man-years to be implemented, while it remains virtually impossible to provide a complete coverage of the language. In particular, in term-intensive documents, where the constituent words focus in a particular area, existing general-purpose morphological lexica fail to provide a sufficient coverage. In this article, a methodology is proposed for the generation of specialised morphological lexica using term-intensive documents. This methodology is based on the matching-and-masking technique, in combination with language-specific constraints. Experimental results indicate that this method generates results of a high accuracy.

1. ΕΙΣΑΓΩΓΗ

Στην μελέτη της Ελληνικής γλώσσας, ιδιαίτερη σημασία έχει η κατασκευή του αντίστοιχου μορφολογικού λεξικού. Η συμβατική μέθοδος κατασκευής λεξικών απαιτεί την καταγραφή των διαφόρων μορφών των λέξεων της γλώσσας, ώστε να εξαχθούν τα

αντίσταιχα λήμματα μαζί με τον τρόπο κλίσης τους. Η καταγραφή αυτή αποτελεί μία επίπονη διαδικασία, η ολοκλήρωση της οποίας χρειάζεται εργασία της τάξεως των ανθρωποετών. Σαν παράδειγμα μπορεί να αναφερθεί το μορφολογικό λεξικό το οποίο έχει κατασκευασθεί στο Ινστιτούτο Επεξεργασίας του Λόγου (ΙΕΛ) [1] και απαίτησε χρονικό διάστημα 10 περίπου ετών για να δημιουργηθεί. Έως σήμερα, το λεξικό αυτό εμπλουτίζεται ώστε να επιτευχθεί η καλύτερη δυνατή κάλυψη της Ελληνικής γλώσσας, με την μεγάλη ποικιλία τύπων που εμφανίζει στην εποχή μας.

Ένα αναπόφευκτο μειονέκτημα της χειρωνακτικής διαδικασίας κατασκευής λεξικών, που επιτείνεται λόγω της δυναμικής εξέλιξης της κάθε γλώσσας, είναι ότι το παραγόμενο μορφολογικό λεξικό αδυνατεί να καλύψει το σύνολο των λέξεων της γλώσσας. Το φαινόμενο αυτό παρουσιάζεται εντονότερο σε γλώσσες που διαθέτουν μεγάλο αριθμό τύπων λέξεων (πολυτυπία), στις οποίες συγκαταλέγεται και η Ελληνική. Κατά συνέπεια, η πλειοψηφία των υπαρχόντων μορφολογικών λεξικών αποσκοπεί να καλύψει τις συχνότερες λέξεις, καθιστώντας τα λεξικά αυτά κατάλληλα για κείμενα που χρησιμοποιούν γενική γλώσσα. Όμως, στην περίπτωση κειμένων εντάσεως όρων, όπου οι χρησιμοποιούμενες λέξεις είναι σε μεγάλο βαθμό εξειδικευμένες σε μία συγκεκριμένη θεματική περιοχή, τα υπάρχοντα μορφολογικά λεξικά δεν παρέχουν επαρκή κάλυψη. Το πρόβλημα αυτό εντείνεται όταν τα υπό μελέτη κείμενα αναφέρονται σε συγκεκριμένους επιστημονικούς τομείς με εξειδικευμένη ορολογία.

Στην παρούσα ανακοίνωση προτείνεται μία μέθοδος η οποία επιτρέπει την κατασκευή μορφολογικών λεξικών εστιασμένων σε μία συγκεκριμένη γνωστική περιοχή. Τα λεξικά αυτά κατασκευάζονται με αυτόματο τρόπο, μέσω ενός αλγορίθμου, και κατά συνέπεια απαιτούν το ελάχιστο δυνατό κόστος σε ανθρώπινους πόρους. Ανάλογα με τα κείμενα που χρησιμοποιούνται σαν πηγή δεδομένων του συστήματος, καθίσταται δυνατή η κατασκευή λεξικών τα οποία είναι εξειδικευμένα ώστε να καλύπτουν μία συγκεκριμένη θεματική περιοχή.

2. ΠΕΡΙΓΡΑΦΗ ΤΟΥ ΣΥΣΤΗΜΑΤΟΣ

Η έρευνα στα θέματα της μορφολογικής επεξεργασίας λεκτικών μονάδων έχει ακολουθήσει δύο κύριες κατευθύνσεις. Η πρώτη κατεύθυνση χαρακτηρίζεται από την χρήση γλωσσολογικών κανόνων (όπως επί παραδείγματι στο άρθρο [3]) για την μοντελοποίηση των μορφολογικών φαινομένων, την συμπίεση της πληροφορίας στα μορφολογικά λεξικά, την λημματοποίηση καθώς και την παραγωγή των μορφολογικών τύπων. Η δεύτερη κατεύθυνση χαρακτηρίζεται από την χρήση υπολογιστικών αρχιτεκτονικών που προσομοιώνουν την δομή βιολογικών νευρωνικών δικτύων, ώστε να επιτευχθεί η μορφολογική επεξεργασία των λέξεων (όπως, επί παραδείγματι στο [2]). Η μεθοδολογία που αποτελεί το αντικείμενο της συγκεκριμένης ανακοίνωσης είναι υβριδικού τύπου, με την έννοια ότι χρησιμοποιεί και

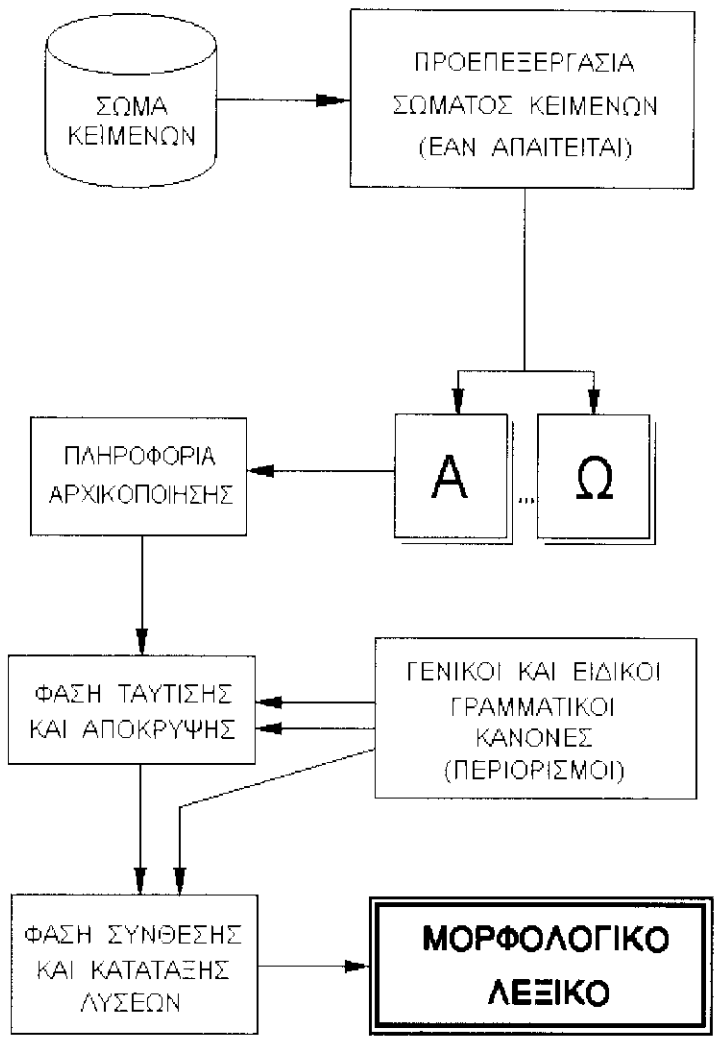
αξιοποιεί αριθμό κανόνων, αλλά, ενσωματώνει επίσης κάποιους στατιστικούς αλγορίθμους, όπως είναι η τάση σήμερα σε πολλές εφαρμογές της γλωσσικής τεχνολογίας. Η δομή του συστήματος παρουσιάζεται στο σχήμα 1.

Η προτεινόμενη μεθοδολογία βασίζεται στην τεχνική της ταύτισης-και-απόκρυψης (matching-and-masking technique), η οποία χρησιμοποιείται ευρέως σε εφαρμογές αναγνώρισης προτύπων (pattern recognition). Σύμφωνα με την τεχνική αυτή, το κάθε πρότυπο διαιρείται σε τμήματα και επιχειρείται η ταύτιση ομοειδών τμημάτων των προτύπων ενώ ταυτόχρονα αποκρύπτονται προσωρινά τα υπόλοιπα τμήματα των προτύπων. Στην συγκεκριμένη περίπτωση, όπου τα πρότυπα είναι λεκτικές μονάδες, η κάλυψη και απόκρυψη αναφέρεται στα θέματα και τις αντίστοιχες καταλήξεις των λεκτικών μονάδων. Έτσι, ας υποτεθεί ότι υπάρχουν δύο συγκεκριμένες λέξεις/πρότυπα οι οποίες παριστάνονται ως $\theta_1\kappa_1$ και $\theta_2\kappa_2$, όπου θi το εκάστοτε θέμα και κi η αντίστοιχη κατάληξη. Τότε, η επιτυχής ταύτιση συνεπάγεται ότι είτε $\theta_1 = \theta_2$ (οπότε ταυτίζονται τα θέματα των δύο λέξεων) είτε $\kappa_1 = \kappa_2$ (οπότε ταυτίζονται οι καταλήξεις των δύο λέξεων).

Η διαδικασία εντοπισμού πιθανών θεμάτων και καταλήξεων υλοποιείται με την επαναληπτική εφαρμογή της τεχνικής ταύτισης-και-απόκρυψης. Σε κάθε επανάληψη, μελετώνται όλες οι λεκτικές μονάδες που υπάρχουν στο σώμα κειμένων ώστε από τις ήδη προσδιορισθείσες καταλήξεις και θέματα να προσδιορισθούν νέες πιθανές καταλήξεις και θέματα. Έτσι, σταδιακά δημιουργείται ένα σύνολο από πιθανές καταλήξεις και θέματα για κάθε λεκτική μονάδα. Πριν την έναρξη της διαδικασίας αυτής, παρέχονται στον αυτόματο μορφολογικό επεξεργαστή ως εκ των προτέρων γνώση (a priori knowledge) 50 καταλήξεις της Ελληνικής γλώσσας, οι οποίες χρησιμοποιούνται για την αρχικοποίηση του συστήματος.

Στο προτεινόμενο σύστημα, η αρχή ταύτισης-και-απόκρυψης συμπληρώνεται από ένα σύνολο γραμματικών περιορισμών ως προς τις αποδεκτές και μη αποδεκτές μορφές των θεμάτων και καταλήξεων στην Ελληνική γλώσσα [4], [5]. Ένα παράδειγμα πιθανού περιορισμού θα ήταν ότι κάθε έγκυρη κατάληξη της Ελληνικής γλώσσας θα πρέπει να περιέχει τουλάχιστον ένα φωνήεν. Οι κανόνες και περιορισμοί αυτοί χρησιμοποιούνται για να περιορίσουν τις πιθανές καταλήξεις και θέματα στις οποίες αναλύεται κάθε δεδομένη λέξη.

Με την παραπάνω διαδικασία, για κάθε λέξη του σώματος παράγεται ένας αριθμός από δυνατές λύσεις διαχωρισμού σε θέμα και κατάληξη. Αυτές οι λύσεις κατατάσσονται ανάλογα με την πιθανότητά τους να αποτελούν τον βέλτιστο διαχωρισμό της δεδομένης λεκτικής μονάδας σε θέμα και κατάληξη, χρησιμοποιώντας ένα κριτήριο κατάταξης. Στα πειράματα που θα περιγραφούν στην επόμενη παράγραφο, σαν κριτήριο κατάταξης χρησιμοποιείται το ελάχιστο μήκος κατάληξης. Η προτεινόμενη μεθοδολογία επιτρέπει, μέσω της επεξεργασίας των λέξεων ενός σώματος κειμένων, την κατασκευή του αντίστοιχου μορφολογικού λεξικού.



Σχήμα 1: Διαγραμματική περιγραφή του συστήματος αυτοματοποιημένης μορφολογικής επεξεργασίας.

3. ΠΕΙΡΑΜΑΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ

Για την μελέτη της συμπεριφοράς του αυτόματου μορφολογικού επεξεργαστή, εξετάσθηκαν τρία σώματα κειμένων τα οποία θα αναφέρονται ως *Σ1, Σ2* και *Σ3*. Το πρώτο σώμα κειμένων, *Σ1*, είναι γενικού τύπου και χρησιμοποιείται για σκοπούς σύγκρισης με τα κείμενα εντάσεως όρων. Έτσι, το *Σ1* περιλαμβάνει λογοτεχνικό κείμενο, το οποίο μπορεί να θεωρηθεί ότι - με την εξαίρεση κάποιων ιδιωματικών κλίσεων και μορφών - χαρακτηρίζεται από μη εξειδικευμένη γλώσσα και άρα θα πρέπει να καλύπτεται σε σχετικά υψηλό βαθμό από γενικής χρήσης μορφολογικά λεξικά. Το μέγεθος του *Σ1* είναι 199.494 λέξεις, επιλεγμένο ώστε να παράγει παρόμοιο πλήθος λεκτικών μορφών με τα δύο σώματα κειμένων εντάσεως όρων, *Σ2* και *Σ3*. Το σώμα *Σ2* αποτελείται από τρία κείμενα σε θέματα αεροναυπηγικής ενώ το σώμα Σ3 αποτελείται από τρία κείμενα σε θέματα βιολογίας. Το σώμα κειμένων *Σ2* περιλαμβάνει 245.006 λέξεις ενώ το σώμα *Σ3* περιλαμβάνει 282.902 λέξεις.

Στο στάδιο προεπεξεργασίας, για κάθε σώμα κειμένων παραλείπονται όλες οι λέξεις που περιέχουν αριθμούς ή γράμματα του Λατινικού αλφαβήτου. Επίσης παραλείπονται άκλιτες λέξεις (από ένα σύνολο 180 λέξεων που παρέχονται στο σύστημα ως εκ των προτέρων γνώση) ενώ για κάθε λέξη του εκάστοτε σώματος που δεν εμπίπτει στους παραπάνω περιορισμούς, διατηρείται μόνο η πρώτη της εμφάνιση. Σαν αποτέλεσμα, δημιουργείται μία λίστα που περιέχει τις λεκτικές μονάδες του σώματος κειμένων που πρόκειται να επεξεργασθεί το σύστημα. Ο αριθμός διαφορετικών λεκτικών μονάδων ανά σώμα κειμένων ανέρχεται σε 20.000 στην περίπτωση του *Σ1*, 13.905 στην περίπτωση του *Σ2* και 20.804 στην περίπτωση του *Σ3*.

Σαν μέτρα αξιολόγησης της προτεινόμενης μεθόδου μορφολογικής επεξεργασίας χρησιμοποιήθηκαν δύο διαφορετικές εκδόσεις του μορφολογικού λεξικού του ΙΕΛ. Η βασική έκδοση (που θα συμβολίζεται με ΛεξΙΕΛ1) περιλαμβάνει περίπου 50.000 διαφορετικά λήμματα της Ελληνικής γλώσσας και συνολικά 92.367 αλλόμορφα. Η επαυξημένη έκδοση (που θα συμβολίζεται με ΛεξΙΕΛ2) περιλαμβάνει πάνω από 63.000 διαφορετικά λήμματα της Ελληνικής γλώσσας και συνολικά 107.026 αλλόμορφα.

Για την μελέτη των πειραματικών αποτελεσμάτων, οι λέξεις που αποτελούν το κάθε κείμενο χωρίζονται σε δύο κατηγορίες:

(i) τις λέξεις που περιέχονται στο εκάστοτε μορφολογικό λεξικό και

(ii) τις λέξεις που δεν περιέχονται στο εκάστοτε μορφολογικό λεξικό.

	Σώμα Σ1	**Σώμα Σ2**	**Σώμα Σ3**
Λεξικό *ΛεξΙΕΛ1*	89,0%	84,2%	74,4%
Λεξικό *ΛεξΙΕΛ2*	90,9%	86,3%	76,3%
Αυτομ. Μορφολογικός Επεξεργαστής	98,9%	98,6%	99,3%

Πίνακας 1: Κάλυψη των λέξεων των σωμάτων κειμένων.

Όπως φαίνεται στον πίνακα 1, στην περίπτωση του σώματος κειμένων *Σ1* περίπου το 11% των λεκτικών μορφών είναι άγνωστες στο βασικό μορφολογικό λεξικό του ΙΕΛ (ΛεξΙΕΛ1), ποσοστό που περιορίζεται στο 9% στο επαυξημένο μορφολογικό λεξικό του ΙΕΛ (ΛεξΙΕΛ2). Στην περίπτωση του σώματος κειμένων *Σ2*, περίπου 16% και 14% των λεκτικών μορφών είναι άγνωστες στο βασικό και στο επαυξημένο μορφολογικό λεξικό του ΙΕΛ αντίστοιχα. Στην περίπτωση του σώματος κειμένων *Σ3*, τα ποσοστά αυτά είναι πολύ υψηλότερα και προσεγγίζουν το 26% και 24% για το απλό και το επαυξημένο μορφολογικό λεξικό αντίστοιχα. Συνεπώς, όπως αναμενόταν, παρατηρείται μεγαλύτερος αριθμός αγνώστων λέξεων στις περιπτώσεις σωμάτων κειμένων με ένταση όρων απ' ότι σε σώματα που χρησιμοποιούν μη εξειδικευμένη γλώσσα.

Οι λεκτικές μονάδες τις οποίες ο αυτόματος μορφολογικός επεξεργαστής δεν διαχωρίζει σε θέμα και κατάληξη ανέρχονται σε περίπου 1,5% στην περίπτωση του σώματος *Σ2* και σε λιγότερο από 1% στην περίπτωση του σώματος *Σ3*. Τα ποσοστά αυτά είναι της ίδιας τάξης μεγέθους όπως για το σώμα κειμένων *Σ1*. Συνεπώς, συγκρίνοντας τον αυτόματο μορφολογικό επεξεργαστή με τα μορφολογικά λεξικά παρατηρείται - όπως αναμενόταν - ότι ο αυτόματος μορφολογικός επεξεργαστής παρέχει μία κάλυψη πολύ υψηλότερη από αυτές των δύο μορφολογικών λεξικών γενικού τύπου.

	1η λύση	2η λύση	3η λύση	4η λύση	Αποτυχία
Σ1 & ΛεξΙΕΛ1	93.7%	4.9%	0.3%	0.0%	1.1%
Σ1 & ΛεξΙΕΛ2	93.7%	4.8%	0.4%	0.0%	1.1%
Σ2 & ΛεξΙΕΛ1	93.9%	4.4%	0.1%	0.0%	1.6%
Σ2 & ΛεξΙΕΛ2	93.8%	4.3%	0.2%	0.0%	1.7%
Σ3 & ΛεξΙΕΛ1	94.0%	4.9%	0.1%	0.0%	1.0%
Σ3 & ΛεξΙΕΛ2	93.9%	5.0%	0.1%	0.0%	1.0%

Πίνακας 2: Συμφωνία περιεχομένων των δύο μορφολογικών λεξικών με τις λύσεις του αυτόματου μορφολογικού επεξεργαστή.

Ενδιαφέρον παρουσιάζει η μελέτη της ακρίβειας των λύσεων που παράγονται από τον αυτόματο μορφολογικό επεξεργαστή. Η σύγκριση των αποτελεσμάτων του αυτόματου μορφολογικού επεξεργαστή ως προς τα περιεχόμενα των δύο μορφολογικών λεξικών παρουσιάζεται στον πίνακα 2, όπου τα αντίστοιχα ποσοστά αναφέρονται στις λέξεις οι οποίες περιέχονται στο εκάστοτε μορφολογικό λεξικό. Παρατηρείται ότι για το 94% περίπου των λεκτικών μονάδων, η πρώτη κατά σειρά λύση που παράγει ο

αυτόματος μορφολογικός επεξεργαστής ταυτίζεται με την αντίστοιχη πληροφορία που περιέχεται στα μορφολογικά λεξικά. Η δεύτερη λύση που παράγει ο αυτόματος μορφολογικός επεξεργαστής ταυτίζεται με την πληροφορία που εμπεριέχεται στο μορφολογικό λεξικό για το 4% έως 5% των λεκτικών μονάδων. Οι περιπτώσεις που η τρίτη κατά σειρά λύση είναι αυτή που βρίσκεται καταχωρημένη στο μορφολογικό λεξικό είναι εξαιρετικά σπάνια, με ποσοστό της τάξης του 0,1%, ενώ δεν παρατηρείται καμμία περίπτωση όπου η καταχωρημένη στο λεξικό λύση να κατατάσσεται στην τέταρτη ή σε χαμηλότερη θέση. Αξιοσημείωτη είναι η σταθερότητα της κατανομής των σωστών μορφολογικών διαχωρισμών στις κατηγορίες πιθανών λύσεων του αυτόματου μορφολογικού επεξεργαστή, για τα τρία διαφορετικά σώματα κειμένων και τις δύο εκδόσεις του μορφολογικού λεξικού.

Τα παραπάνω πιστοποιούν ότι το προτεινόμενο σύστημα αυτόματης μορφολογικής επεξεργασίας παράγει την γραμματικά σωστή λύση (όπως ορίζεται από το μορφολογικό λεξικό) για την πλειοψηφία των λεκτικών μονάδων που συναντώνται στα σώματα κειμένων. Αυτό επιτυγχάνεται χρησιμοποιώντας έναν αυτοματοποιημένο αλγόριθμο επεξεργασίας στον οποίο παρέχεται αρχικά μία πολύ μικρή ποσότητα γραμματικής/ γλωσσολογικής πληροφορίας. Κατά συνέπεια, η προτεινόμενη μέθοδος μπορεί να αποτελέσει μία αξιόπιστη μεθοδολογία για την αυτόματη δημιουργία μορφολογικών λεξικών με ορολογικό πλούτο βάσει κειμένων εντάσεως όρων.

4. ΣΥΜΠΕΡΑΣΜΑΤΑ

Στην ανακοίνωση αυτή, παρουσιάσθηκε ένα σύστημα για την αυτόματη δημιουργία μορφολογικών λεξικών της Ελληνικής γλώσσας, τα οποία κατασκευάζονται με βάση κείμενα ορολογικού πλούτου. Το σύστημα αυτό βασίζεται στην τεχνική "κάλυψης-και-απόκρυψης", με χρήση a priori γνώσης σχετικής με την Ελληνική γλώσσα. Πειραματικά αποτελέσματα έδειξαν ότι το σύστημα αυτό επιτρέπει την επεξεργασία λέξεων από κείμενα με ορολογικό πλούτο, έχοντας ως αποτέλεσμα μορφολογικά λεξικά τα οποία είναι εξειδικευμένα ώστε να καλύπτουν τις αντίστοιχες γνωστικές περιοχές. Η προτεινόμενη μέθοδος μπορεί να χρησιμοποιηθεί και για αυτόματο εμπλουτισμό έτοιμων μορφολογικών λεξικών με βάση κείμενα εντάσεως όρων. Το αντικείμενο αυτό αποτελεί και το θέμα της μελλοντικής έρευνας στο ΙΕΛ γύρω από τις τεχνικές αυτές.

5. ΑΝΑΦΟΡΕΣ

[1] Αγλαμίσης, Ι. & Μαντζαρη, Ε. (1995) Μορφολογικό Λεξικό - Γραμματικός Χαρακτηριστής. Κείμενο Εργασίας ELL/STD/DEPT/9501, Ινστιταύτο Επεξεργασίας του Λόγου, Αθήνα.

[2] Gasser, M. (1994) Modularity in a Connectionist Model of Morphology Acquisition. In Proceedings of the 15th International Conference on Computational Linguistics, August 5-9, 1994, Kyoto, Japan, Vol. 1, pp. 214-220.

[3] Pentheroudakis, J. & Vanderwende, L. (1993) Automatically Identifying Morphological Relations in Machine-Readable Dictionaries. Technical Report MSR-TR-93-06, Microsoft Research Advanced Technology Division, Microsoft Corporation, One Microsoft Way, Redmond, WA., 98052.

[4] Ράλλη, Α. (1986) Κλίση και Παραγωγή. Πρακτικά της 7ης Ετήσιας Συνάντησης του Τομέα Γλωσσολογίας της Φιλοσοφικής Σχολής του Αριστοτέλειου Πανεπιστημίου Θεσσαλονίκης, 12-14 Μαΐου 1986, pp. 29-48.

[5] Τριανταφυλλίδης, Μ. (1941) Νέα Ελληνική Γραμματική (Δημοτική). Ανάτυπο με διορθώσεις 1978. Ινστιτούτο Μοντέρνων Ελληνικών Σπουδών, Θεσσαλονίκη.

Ταμπουρατζής Γεώργιος, Καραγιάννης Γεώργιος,

Ινστιτούτο Επεξεργασίας του Λόγου,

Αρτέμιδος 6 και Επιδαύρου,

Παράδεισος Αμπρουσίου, 151 25.