

ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΚΟΙΝΩΝΙΑ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ

ΠΡΑΞΗ

«ΕΠΕΞΕΡΓΑΣΙΑ ΕΙΚΟΝΩΝ, ΗΧΟΥ ΚΑΙ ΓΛΩΣΣΑΣ»

στο πλαίσιο του ΜΕΤΡΟΥ 3.3

«Έρευνα και Τεχνολογική Ανάπτυξη στην Κοινωνία της Πληροφορίας»

ΕΡΓΟ - 9: ΙΑΤΡΟΛΕΞΗ

ΠΑΡΑΔΟΤΕΟ

Π6: ΚΑΝΟΝΕΣ ΑΝΑΓΝΩΡΙΣΗΣ ΠΟΛΥΛΕΚΤΙΚΩΝ ΌΡΩΝ ΒΙΟΪΑΤΡΙΚΗΣ

(στο παραδοτέο περιλαμβάνεται οπτικός δίσκος που περιέχει τους κανόνες)

Ημερομηνία:	28.02.2007
Έκδοση:	Final
Τύπος:	Εμπιστευτικό

ΠΕΡΙΕΧΟΜΕΝΑ

1. ΕΙΣΑΓΩΓΗ.....	3
2. ΒΑΣΙΚΕΣ ΟΡΟΛΟΓΙΚΕΣ ΈΝΝΟΙΕΣ	5
3. ΟΡΟΣΤΟΙΧΕΙΑ ΚΑΙ ΣΥΝΘΕΤΙΚΑ ΤΟΥ ΠΟΛΥΛΕΚΤΙΚΟΎ ΣΥΜΠΛΟΚΟΥ ΌΡΟΥ	7
4. ΟΡΟΛΟΓΙΚΗ ΑΝΑΛΥΣΗ ΤΟΥ ΠΟΛΥΛΕΚΤΙΚΟΥ ΣΥΜΠΛΟΚΟΥ ΌΡΟΥ ΒΙΟΪΑΤΡΙΚΗΣ	10
4.1 Το υλικό της ανάλυσης	10
4.2 Τρόποι σύνθεσης των πολυλεκτικών σύμπλοκων όρων που μελετήθηκαν.....	11
5. ΚΑΝΟΝΕΣ ΑΝΑΓΝΩΡΙΣΗΣ ΠΟΛΥΛΕΚΤΙΚΩΝ ΣΥΜΠΛΟΚΩΝ ΟΡΩΝ ΒΙΟΪΑΤΡΙΚΗΣ	19
5.1 Γενικοί κανόνες σχηματισμού Ονοματικών φράσεων (ΟΦ).....	19
5.2 Γενικοί κανόνες σχηματισμού προθετικών φράσεων (ΠΦ)	20
5.3 Κανόνες αναγνώρισης πολυλεκτικών σύμπλοκων όρων βιοϊατρικής.....	21
5.3.1 Δίλεκτικοί όροι.....	21
5.3.2 Τρίλεκτικοί όροι.....	22
5.3.3 Τετραλεκτικοί όροι.....	24
6. ΟΔΗΓΙΕΣ ΧΡΉΣΗΣ ΤΟΥ ΟΠΤΙΚΟΎ ΔΪΣΚΟΥ.....	27
7. ΒΙΒΛΙΟΓΡΑΦΪΑ.....	28

1. ΕΙΣΑΓΩΓΗ

Η παρούσα αναφορά, που αποτελεί το παραδοτέο Π6 του έργου, εντάσσεται στην Ενότητα Εργασίας 3 και περιγράφει αναλυτικά τους «**Κανόνες αναγνώρισης των πολυλεκτικών όρων βιοϊατρικής**». Στόχο της ΕΕ3 αποτελεί η επεξεργασία των συλλεχθέντων βιοϊατρικών κειμένων και άλλων λεξικογραφικών πηγών στο πεδίο της βιοϊατρικής προκειμένου να εξαχθεί η ορολογία που θα αποτελέσει τη βάση πάνω στην οποία θα χτιστεί η οντολογία του ΙΑΤΡΟΛΕΞΗ.

Η συγκέντρωση των **μονολεκτικών όρων** στηρίχθηκε, εκτός από την **επεξεργασία υπαρχόντων λεξιλογίων** στον τομέα της βιοϊατρικής, στην εξαγωγή όρων από τα κείμενα, αξιοποιώντας εν πολλοίς το μεγάλο ποσοστό κάλυψης του λημματολογίου του Μορφολογικού Λεξικού της Neurosoft και αναγνωρίζοντας ως όρους «άγνωστες λέξεις» που απαντούν με μεγάλη συχνότητα στο Σώμα Ελληνικών Βιοϊατρικών Κειμένων (ΣΕΒΚ) του ΙΑΤΡΟΛΕΞΗ (βλ. Παραδοτέο 5 «Μονολεκτικοί όροι βιοϊατρικής»).

Η διαδικασία συγκέντρωσης των **πολυλεκτικών όρων** που θα εμπλουτίσουν την οντολογία αποτελεί σαφώς συνθετότερη και πολυπλοκότερη ορολογική και γλωσσολογική εργασία. Για την **εξαγωγή πολυλεκτικών όρων** από κείμενα απαιτείται η διατύπωση **μιας τυπικής γραμματικής σχηματισμού των πολυλεκτικών όρων** η οποία θα τροφοδοτεί τον Αναγνωριστή Όρων, προϊόν της Ενότητας Εργασίας 2 του έργου.

Η παρούσα αναφορά αποτελεί την αναλυτική περιγραφή των **κανόνων σχηματισμού των πολυλεκτικών όρων**, όπως αυτοί προέκυψαν μετά από μελέτη και αναλυτική καταγραφή των δομών σχηματισμού περίπου 5.000 πολυλεκτικών όρων βιοϊατρικής.

Πιο αναλυτικά, η παρούσα αναφορά περιλαμβάνει τις ακόλουθες ενότητες:

- **Ενότητα 2 – Βασικές ορολογικές έννοιες:** Παρουσίαση των βασικών εννοιών και ορισμών σχετικά με τα είδη των όρων και τη μορφή τους (εσωτερική και εξωτερική).
- **Ενότητα 3 – Οροστοιχεία και συνθετικά του πολυλεκτικού σύμπλοκου όρου:** Εξήγηση του ρόλου των άμεσων συνθετικών (προσδιοριζόμενου και προσδιοριστικού συνθετικού) και του λεκτικού μήκους του πολυλεκτικού σύμπλοκου όρου.
- **Ενότητα 4 – Ορολογική ανάλυση του πολυλεκτικού σύμπλοκου όρου βιοϊατρικής:** Αναλυτική καταγραφή των τρόπων και των δομών σύνθεσης των πολυλεκτικών όρων με βάση σώμα υλικού περίπου 5.000 όρων βιοϊατρικής.
- **Ενότητα 5 – Κανόνες αναγνώρισης των πολυλεκτικών όρων βιοϊατρικής:** Διατύπωση κανόνων για το σχηματισμό απλών και σύνθετων

ονοματικών φράσεων με στόχο την αναγνώριση και εξαγωγή από κείμενα βιοϊατρικής πολυλεκτικών σύμπλοκων όρων του πεδίου.

- **Ενότητα 6 – Οδηγίες χρήσης του οπτικού δίσκου:** Περιγράφει τον τρόπο χρήσης του επισυναπτόμενου οπτικού δίσκου, ο οποίος περιέχει τους κανόνες αναγνώρισης.

2. ΒΑΣΙΚΕΣ ΟΡΟΛΟΓΙΚΕΣ ΕΝΝΟΙΕΣ

Με βάση το Πρότυπο ISO 1087 -1 και την ελληνική έκδοσή του, ΕΛΟΤ 561-1:

Οι έννοιες (concepts) είναι νοητικοί αντιπρόσωποι μεμονωμένων αντικειμένων (objects) ή συνόλων αντικειμένων με κοινά χαρακτηριστικά. Οι έννοιες αποτελούν βασικές μονάδες σκέψης που συγκροτούνται με αφαίρεση βάσει ορισμένων ιδιοτήτων τους. Κάθε έννοια μπορεί να περιγραφεί γλωσσικά με μια πρόταση, τον ορισμό της (definition), και να προσδιοριστεί με μια γλωσσική έκφραση, τον αντίστοιχο όρο (term).

Οι όροι διακρίνονται σε **απλούς** και **σύμπλοκους**. Ο **απλός όρος** αποτελείται από ένα θέμα (stem) με ή χωρίς πρόσθετα προθήματα (prefixes) και με ή χωρίς κατάληξη (ending). Οι απλοί όροι είναι μονοθεματικές λέξεις (δηλαδή ριζικές ή παράγωγες λέξεις). Ο **σύμπλοκος όρος** αποτελείται από περισσότερα από ένα θέματα με ή χωρίς πρόσθετα προθήματα ή επιθήματα και με ή χωρίς καταλήξεις. Οι σύμπλοκοι όροι είναι πολυθεματικές λέξεις (δηλαδή σύνθετες ή παρασύνθετες λέξεις) ή φράσεις. Οι σύμπλοκοι όροι διακρίνονται περαιτέρω σε μονολεκτικούς ή πολυλεκτικούς¹. Παραδείγματα:

απλοί όροι (simple terms):

νευρ-ίτ-ίδα
νεύρ-ωσ-η
πνευμον-ία
 κτλ.

σύμπλοκοι όροι (complex terms):

μονολεκτικοί
πνευμ-ο-θώρακ-ας
πυελ-ο-νεφρ-ίτιδα
υπερ-παρα-θυρεοειδ-ισμός
 κτλ.

ΠΟΛΥΛΕΚΤΙΚΟΙ:

εμβόλιο AIDS
ιός ηπατίτιδας A
τρανσφεράση νουκλεοτιδυλομάδας DNA
πρωτεΐνη εξωτερικής μεμβράνης βακτηρίων

¹ Ο πολυλεκτικός σύμπλοκος όρος λέγεται και **σύνθετος όρος** (compound term).

Η εξωτερική μορφή ενός όρου είναι η διάταξη των φθόγγων ή φωνημάτων ή αντίστοιχων προς αυτά γραμμάτων, που αποτελούν τον όρο. Η **εξωτερική μορφή** ενός όρου όπως αρθρώνεται με φθόγγους ή φωνήματα λέγεται **φωνητική μορφή**, ενώ η εξωτερική μορφή του όρου όπως παριστάνεται με γράμματα λέγεται **γραπτή μορφή** του όρου.

Η **εσωτερική μορφή** του απλού όρου ταυτίζεται με τη σημασία του, ενώ η εσωτερική μορφή του σύμπλοκου πολυλεκτικού όρου είναι η παράθεση των κύριων σημασιών των συστατικών του.

Η συντόμευση της **εξωτερικής μορφής (σύντμηση)** ενός πλήρους όρου με παράλειψη οποιουδήποτε μέρους του, χωρίς μεταβολή στην έννοια που ο πλήρης όρος αντιπροσωπεύει, οδηγεί στο συντετμημένο όρο (abbreviated term), που είναι ταυτόσημος με τον πλήρη όρο.

Αν παραλείπονται κάποια γράμματα ενός πλήρους απλού όρου, τότε έχουμε μια συντομομορφή (abbreviation). Το **αρκτικόλεξο** (initialism) είναι **συντετμημένος σύμπλοκος όρος**, στον οποίο διατηρούνται τα αρχικά γράμματα των λέξεων του πλήρους σύμπλοκου όρου, και ο οποίος προφέρεται ή συλλαβικά ή ως μια λέξη ή με χωριστά γράμματα ή και με τους δύο τρόπους:

ΑΕ = Ανώνυμη Εταιρεία
ΙΧ = Ιδιωτικής Χρήσης
DNA = Deoxyribonucleic Acid

Το **ακρώνυμο** (acronym) είναι **συντετμημένος σύμπλοκος όρος**, ο οποίος χρησιμοποιείται ως σειρά γραμμάτων που λαμβάνονται από τις λέξεις του πλήρους σύμπλοκου όρου και προφέρεται πάντοτε συλλαβικά ως μια λέξη:

ΟΤΕ (Οργανισμός Τηλεπικοινωνιών Ελλάδος)
 ΑΣΕΠ (Ανώτατο Συμβούλιο Επιλογής Προσωπικού)

3. ΟΡΟΣΤΟΙΧΕΙΑ ΚΑΙ ΣΥΝΘΕΤΙΚΑ ΤΟΥ ΠΟΛΥΛΕΚΤΙΚΟΥ ΣΥΜΠΛΟΚΟΥ ΟΡΟΥ

Στην ενότητα που ακολουθεί δίνονται μια σειρά **όρων** και **ορισμών** που θα χρησιμοποιηθούν κατά την περιγραφή των τρόπων σχηματισμού των πολυλεκτικών σύμπλοκων όρων και θα αξιοποιηθούν στη διατύπωση των κανόνων σχηματισμού τους.

ορολογική ανάλυση

ανάλυση ενός **όρου** στα **οροσυνθετικά** του

- σημείωση 1 – Στην ορολογική ανάλυση προεξάρχει η *εννοιολογική διάσταση*, σε αντίθεση με τη γραμματική ανάλυση όπου προεξάρχει η *μορφολογική διάσταση*.
- σημείωση 2 – Ο **απλός όρος** αποτελείται από ένα **οροσυνθετικό**.
- σημείωση 3 – Ο **σύμπλοκος όρος**, σε πρώτη φάση ορολογικής ανάλυσης, χωρίζεται – κατά κανόνα – σε δύο οροσυνθετικά, τα λεγόμενα **άμεσα συνθετικά**.

οροστοιχείο

μορφολογικό μέρος ενός **όρου** το οποίο είναι φορέας σημασίας

- σημείωση – Το οροστοιχείο μπορεί να είναι **λέξη**, **φράση**, ή **μόρφημα**.
Παράδειγμα: **συν/τελ/εσ/τ/ής / ηχ/ο/από/ρ/ρόφ/ησ/ης**

λέξη

οροστοιχείο που μπορεί να υπάρξει ως αυθύπαρκτη χωριστή μονάδα σε μια πρόταση

- σημείωση – Στο γραπτό λόγο η λέξη είναι – συνήθως – μέρος του όρου που βρίσκεται μεταξύ δύο διαδοχικών κενών διαστημάτων.
Παράδειγμα: **συντελεστής/ηχοαπρρόφησης**.

φράση

συντεταγμένο σύνολο **λέξεων** που αποτελεί από μόνο του έναν **όρο** ή έναν προσδιορισμό

- σημείωση – Παράδειγμα: **μέσος / συντελεστής ηχοαπρρόφησης**.

οροσυνθετικό

μέρος **όρου** που από μόνο του αποτελεί **όρο** ή υποδηλώνει **χαρακτηριστικό** στο πλαίσιο του υποκείμενου **συστήματος εννοιών**

- σημείωση – Το οροσυνθετικό μπορεί να αποτελείται από ένα ή περισσότερα οροστοιχεία.

άμεσο συνθετικό

ένα από τα δύο – συνήθως – **οροσυνθετικά** στα οποία μπορεί να χωριστεί ένας **σύμπλοκος όρος** σε πρώτη φάση **ορολογικής ανάλυσης**

- σημείωση – Κατά κανόνα, το ένα από τα άμεσα συνθετικά είναι το **προσδιοριζόμενο συνθετικό**, ενώ το άλλο είναι το **προσδιορίζον** ή **προσδιοριστικό συνθετικό**.

προσδιοριζόμενο συνθετικό

άμεσο συνθετικό σύμπλοκου όρου που εκφράζει ένα **γένος** της **έννοιας** που υποδηλώνει ο σύμπλοκος όρος

- σημείωση 1 – Στα **παρατακτικά σύνθετα** δεν υπάρχει **προσδιοριζόμενο συνθετικό** και **προσδιορίζον συνθετικό**: τα δύο συνθετικά υπεισέρχονται ισοδύναμα. π.χ. *πομποδέκτης* (= ταυτόχρονα *πομπός* και *δέκτης*)
- σημείωση 2 – Παράδειγμα προσδιοριζόμενου συνθετικού: **αντίσωμα** ηπατίτιδας.

προσδιορίζον συνθετικό, προσδιοριστικό συνθετικό

άμεσο συνθετικό σύμπλοκου όρου που εκφράζει το **διακριτικό χαρακτηριστικό** το οποίο μετατρέπει το **γένος** που υποδηλώνει το προσδιοριζόμενο συνθετικό σε **είδος**

- σημείωση 1 – Στα **παρατακτικά σύνθετα** δεν υπάρχει **προσδιοριζόμενο συνθετικό** και **προσδιορίζον συνθετικό**: τα δύο συνθετικά υπεισέρχονται ισοδύναμα.
- σημείωση 2 – Παραδείγματα προσδιορίζοντος συνθετικού: **τριφωσφορική** αδενοσίνη, **γενετική** μη διάζευξη

ισοδύναμη γραμματική κατηγορία

γραμματική κατηγορία στην οποία αντιστοιχεί **εννοιολογικά** ο **όρος**

- σημείωση 1 – Ως γραμματικές κατηγορίες εδώ εννοούμε τα γνωστά «μέρη του λόγου» της Γραμματικής.
- σημείωση 2 – Η ισοδύναμη γραμματική κατηγορία ενός οποιουδήποτε **όρου** προκύπτει αμέσως αν απαντήσουμε στην ερώτηση: «**Αν παρίσταναν την έννοια που αντιπροσωπεύει ο όρος μόνο με μία λέξη «X» τι μέρος του λόγου θα ήταν η λέξη «X»;**»
- σημείωση 3 – Παραδείγματα:
 - όργανο του Corti** (Ουσιαστικό)
 - (φροντίδα) **κατ' οίκον** (Επίθετο)
 - (φάρμακο) **τοπικής χρήσης** (Επίθετο)

νιλεκτικότητα, λεκτικό μήκος

ο αριθμός των **λέξεων** από τις οποίες αποτελείται ο **όρος**

- σημείωση 1 – Αν συμβολίσουμε τη **νιλεκτικότητα** με **N**, οι μονολεκτικοί όροι έχουν **N=1**, οι διλεκτικοί όροι έχουν **N=2** κ.ο.κ.

- σημείωση 2 – Παραδείγματα:
 - οξικό βακτηρίδιο **N=2**
 - αιμοποιητικά μητρικά κύτταρα **N=3**
 - φαρμακευτικά εξοπλισμένες ενδομητρικές συσκευές **N=4**
 - κτλ.

ορολογική κατηγορία

κατηγορία στην οποία υπάγεται ο όρος με βάση την διαδοχή των **άμεσων συνθετικών** του

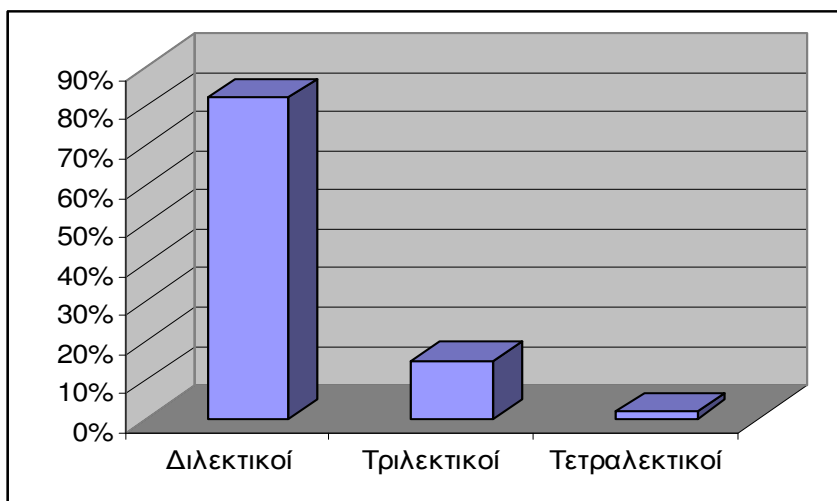
- σημείωση – Η ορολογική κατηγορία μπορεί να είναι:
 - απλός όρος (**ΑΟ**)
 - σύμπλοκος όρος με μη χαρακτηρισίσιμα άμεσα συνθετικά (**ΑΧ**)
 - σύμπλοκος όρος με τη διαδοχή: **προσδιοριστικό-Προσδιοριζόμενο (πΠ)**
 - σύμπλοκος όρος με τη διαδοχή: **Προσδιοριζόμενο- προσδιοριστικό (Ππ)**

4. ΟΡΟΛΟΓΙΚΗ ΑΝΑΛΥΣΗ ΤΟΥ ΠΟΛΥΛΕΚΤΙΚΟΥ ΣΥΜΠΛΟΚΟΥ ΟΡΟΥ ΒΙΟΪΑΤΡΙΚΗΣ

4.1 ΤΟ ΥΛΙΚΟ ΤΗΣ ΑΝΑΛΥΣΗΣ

Η περιγραφή των τρόπων σχηματισμού του πολυλεκτικού σύμπλοκου όρου της βιοϊατρικής στηρίχθηκε στην ορολογική ανάλυση των σχετικών δομών περίπου 5.000 όρων. Η αρχική ανάλυση απέδωσε περίπου 213 διαφορετικές δομές που περιέγραφαν το σύνολο των όρων. Ο χωρισμός του υλικού σε δομές που αφορούν διλεκτικούς, τριλεκτικούς και τετραλεκτικούς όρους επέτρεψε:

1. την ακριβή μέτρηση του αριθμού των αντίστοιχων όρων και την ποσοστιαία συμμετοχή τους στο σύνολο του εξεταζόμενου υλικού. Οι μετρήσεις συνοψίζονται ως εξής:



- Το **82%** των πολυλεκτικών όρων είναι *διλεκτικοί*
 - Το **15%** *τριλεκτικοί*
 - Το **1,9%** *τετραλεκτικοί*
- (το υπόλοιπο ~1% είναι πενταλεκτικοί όροι αμφίβολης εγκυρότητας)

Τα συγκεκριμένα αποτελέσματα μπορούν να αξιοποιηθούν στις στατιστικές μετρήσεις που θα εφαρμοστούν κατά την αναγνώριση και την εξαγωγή των πολυλεκτικών όρων από τα κείμενα.

2. τη μελέτη των πρωτογενών δομών, οι οποίες εμφανίζονται κυρίως στα δίλεκτα, καθώς και των επαυξημένων δομών τους που εμφανίζονται στα τρίλεκτα και τετράλεκτα.

4.2 ΤΡΟΠΟΙ ΣΥΝΘΕΣΗΣ ΤΩΝ ΠΟΛΥΛΕΚΤΙΚΩΝ ΣΥΜΠΛΟΚΩΝ ΟΡΩΝ ΠΟΥ ΜΕΛΕΤΗΘΗΚΑΝ

Στη συνέχεια αυτής της ενότητας παρουσιάζονται πίνακες με τη σύνθεση των διλεκτικών, τριλεκτικών και τετραλεκτικών όρων, από γραμματική, συντακτική και ορολογική άποψη. Στους πίνακες αυτούς, για λόγους οικονομίας, χρησιμοποιούνται στην περιγραφή οι ακόλουθες συντομογραφίες:

Ο	ουσιαστικό
Ο2	ουσιαστικό σε γενική
Ο3	ουσιαστικό σε αιτιατική
Ε	επίθετο
Ε2	επίθετο σε γενική
Ε3	επίθετο σε αιτιατική
ΕΠΡ	επίρρημα
ΜΤΧ	μετοχή
ΠΡΟΘ	πρόθεση: «με», «σε» κτλ.
ΑΡΘΡ	άρθρο
Σ	σύμβολο
Σ-Ο	σύμβολο-ουσιαστικό, π.χ. ν-οξειδίο
Α	αριθμός
Α-Ο	αριθμός-ουσιαστικό, π.χ. 16-διμεθυλοπρωσταγλανδίνη
ΞΛ	ξένη λέξη
ΞΑΚΡ	ξένο αρκτικόλεξο/ακρώνυμο
ΜΟΡ	μόριο «μη»

Οι πίνακες 1Α και 1Β καλύπτουν όλες σχεδόν τις περιπτώσεις διλεκτικών όρων που καταγράφηκαν στο υλικό μας. Στην πρώτη και τέταρτη στήλη περιγράφεται η γραμματική κατηγορία του 1ου και του 2ου συστατικού, ενώ στη δεύτερη και στην τρίτη στήλη καταγράφονται η πρόθεση και το άρθρο, που μεσολαβούν ανάμεσα στα δύο συστατικά. Οι συγκεκριμένες λέξεις δεν καταμετρήθηκαν ως διαφορετικά συστατικά, αφού τόσο η πρόθεση που εισάγει το 2ο συστατικό όσο και το άρθρο αποτελούν μέρη του 2ου συστατικού και όχι ανεξάρτητα οροστοιχεία. Στην πέμπτη στήλη δίνεται χαρακτηριστικό παράδειγμα του περιγραφόμενου τρόπου σύνθεσης, στην έκτη στήλη η συχνότητα με την οποία εμφανίστηκε στο υλικό μας η περιγραφόμενη δομή και στην 7 στήλη περιγράφονται η ισοδύναμη γραμματική κατηγορία και η ορολογική κατηγορία των άμεσων συνθετικών – **Προσδιοριζόμενου** και **προσδιοριστικού** – με τη σειρά που εμφανίζονται στο διλεκτικό όρο.

Στον πίνακα 1Β περιγράφονται χωριστά οι προθετικές φράσεις οι οποίες, αν και αποτελούνται από επαυξημένες ονοματικές φράσεις και θα μπορούσαν να καταμετρηθούν και ως τριλεκτικοί όροι, θεωρήθηκαν παρ' όλα αυτά ως δεύτερα συστατικά διλεκτικών όρων.

Πίνακας 1Α – Τρόποι σύνθεσης διλεκτικών όρων (διλέκτων) βιοϊατρικής

1	2	3	4	5	6	7
1° συστ.	Πρόθεση	Άρθρο	2° συστ.	Παράδειγμα	Συχν.	Ορολ. Κατηγ.
Ο			Ο	βάκιλος μεγαθήριο	10	ΟΥΣ_Ππ
Ο			Σ	βακτηριοφάγος ΜU	193	ΟΥΣ_Ππ
A-O			ΣΑ	16-διμεθυλοπροσταγλανδίνη E2	1	ΟΥΣ_Ππ
Ο			ΞΛ	αδένας Harderian	126	ΟΥΣ_Ππ
Ο		ΑΡΘ	ΞΛ	όργανο του Corti	1	ΟΥΣ_Ππ
Ο			A	ημιχολίνιο 3	13	ΟΥΣ_Ππ
Ο			ΑΣ	κύτταρα 3T3	1	ΟΥΣ_Ππ
Ο			O2	ιός λύσσας	1145	ΟΥΣ_Ππ
Ο		ΑΡΘ	O2	καλύπτρα του μεσεγκεφάλου	18	ΟΥΣ_Ππ
A-O			O2	0-δεμεθυλάση νιτροανισόλης	1	ΟΥΣ_Ππ
A-Σ-O			O2	5-άλφα-ρεδουκτάση τεστοστερόνης	1	ΟΥΣ_Ππ
Ο			Σ-O2	συνθάση N-ακετυλολακτοζαμίνης	5	ΟΥΣ_Ππ
O-Σ			O2	τριφθορομεθοξυφαινυλυδραζόν η-P καρβονυλοκυανιδίου	4	ΟΥΣ_Ππ
Ο			A-O2	δεϋδρογονάσες 17-υδροξυστεροειδών	7	ΟΥΣ_Ππ
Ο		ΑΡΘ	ΞΑΚΡ	πρόδρομοι του RNA	2	ΟΥΣ_Ππ
Ο			E	πρωτέας θαυμαστός	3	ΟΥΣ_Ππ
Ο		ΑΡΘ	E	φύνος ο θαλασσινός	2	ΟΥΣ_Ππ
E			O	ογκογόνοι ιοί	1867	ΟΥΣ_πΠ
E			Σ-O	κυκλικά Ρ-οξίδια	5	ΟΥΣ_πΠ
E			A-O	οξική 17-μεδροξυπρογεστερόνη	1	ΟΥΣ_πΠ
E			ΞΑΚΡ	νεοπλασματικό DNA	20	ΟΥΣ_πΠ
Σ-E			O	ν-δεμεθυλιωτικές οξειδορεδουκτάσες	2	ΟΥΣ_πΠ
A-E			O	2-αμινοαδιπικό οξύ	14	ΟΥΣ_πΠ
A-Σ			O	H-2 αντιγόνα	4	ΟΥΣ_πΠ
Ο	από	ΑΡΘ	O	αντισυλληπτικά από το στόμα	7	ΟΥΣ_Ππ
Ο	από		O	χίμαιρα από ακτινοβολία	22	ΟΥΣ_Ππ
Ο	για	ΑΡΘ	O	αγωγή για το σεξ	1	ΟΥΣ_Ππ
Ο	κατά	ΑΡΘ	O	φροντίδα κατά τη νύχτα	1	ΟΥΣ_Ππ
Ο	κατά	ΑΡΘ		φάρμακα κατά της σχιστοσωμίας	8	ΟΥΣ_Ππ
Ο	κατά		ΞΛ	αναστόμωση κατά Roux-en-Y	1	ΟΥΣ_Ππ
Ο	κατά		O	θεραπεία κατά οικογένεια	3	ΟΥΣ_Ππ
Ο	με	ΑΡΘ	O	θεραπεία με το χορό	1	ΟΥΣ_Ππ
Ο	με		ΞΛ	αγγειοπλαστική με laser	3	ΟΥΣ_Ππ
Ο	με		O	αγγειοπλαστική με μπαλόνι	19	ΟΥΣ_Ππ

Ε.Π. «Κ.τ.Π.» Μέτρο 3.3 «Έρευνα και Τεχνολογική Ανάπτυξη στην ΚτΠ» Πράξη «Επεξεργασία εικόνων, ήχου και γλώσσας»	Έργο 9: «ΙΑΤΡΟΛΕΞΗ»
--	------------------------

Ο	μέσω	ΑΡΘ	Ο	θεραπεία μέσω της μορφής	2	ΟΥΣ_Ππ
Ο	μεταξύ		Ο	δανεισμός μεταξύ βιβλιοθηκών	1	ΟΥΣ_Ππ
Ο	σε	ΑΡΘ	Ο	αντίσταση στα εντομοκτόνα	18	ΟΥΣ_Ππ
Ο	σε		Ο	ευπάθεια σε νόσο	4	ΟΥΣ_Ππ
Ο	χωρίς		Ο	επικοινωνία χωρίς λόγια	1	ΟΥΣ_Ππ
ΣΥΝΟΛΟ					3.538	

Πίνακας 1B – Προθετικές φράσεις ως συστατικά διλεκτικών όρων βιοϊατρικής

1	2						3	4	5
1 ^ο συστ.	2 ^ο συστατικό = Προθετική φράση						Παράδειγμα	Συχν.	Ορολ. Κατηγ.
Ο	κατά	ΑΡΘ		Ο	ΑΡΘ	Ο2	φροντίδα κατά τη διάρκεια της ημέρας	1	Ουσ_Ππ
Ο	με	ΑΡΘ		Ο		Ο2	ακτινοθεραπεία με τη βοήθεια υπολογιστή	1	Ουσ_Ππ
Ο	σε	ΑΡΘ		Ο		Ο2	έρευνα στις υπηρεσίες υγείας	1	Ουσ_Ππ
Ο	από	ΑΡΘ		Ο		Ο2	ικανοποίηση από τη θέση εργασίας	1	Ουσ_Ππ
Ο	με			Ο	ΑΡΘ	Ο2	δηλητηρίαση με μονοξειδίο του άνθρακα	1	Ουσ_Ππ
Ο	με			Ο		Ο2	θεραπεία με εισπνοή οξυγόνου	3	Ουσ_Ππ
Ο	κατά		Ε	Ο			νεοπλάσματα κατά ιστολογικούς τύπους	1	Ουσ_Ππ
Ο	με		Ε	Ο			θεραπεία με φυσικά μέσα	1	Ουσ_Ππ
Ο	με			Ο	και	Ο	αγγειοπλαστική με μπαλόνι και laser	1	Ουσ_Ππ
Ο	σε		Ε	Ο			παροχέτευση σε ορισμένη θέση	1	Ουσ_Ππ
Ο	για		Ε	Ο			ασφάλιση για νοσοκομειακή περίθαλψη	2	Ουσ_Ππ
Ο	από		Ε	Ο			νοσηλεία από αποκλειστική νοσοκόμο	8	Ουσ_Ππ
Ο	από			Ο		Ο2	διαταραχές από χρήση καπνού	4	Ουσ_Ππ

Ε.Π. «Κ.τ.Π.» Μέτρο 3.3 «Έρευνα και Τεχνολογική Ανάπτυξη στην ΚτΠ» Πράξη «Επεξεργασία εικόνων, ήχου και γλώσσας»	Έργο 9: «ΙΑΤΡΟΛΕΞΗ»
--	------------------------

Οι πίνακες 2Α, 2Β και 2Γ καλύπτουν όλες σχεδόν τις περιπτώσεις των τριλεκτικών όρων που καταγράφηκαν στο υλικό μας. Πιο συγκεκριμένα, στον πίνακα 2Α δίνονται οι τρόποι σύνθεσης των τριλέκτων με δομές που αναπτύσσονται δεξιά του προσδιοριζόμενου συνθετικού, το οποίο είναι μια ονοματική φράση. Στον πίνακα 2Β δίνονται οι τρόποι σύνθεσης των τριλέκτων με δομές που αναπτύσσονται αριστερά του προσδιοριζόμενου συνθετικού, το οποίο είναι μια ονοματική φράση, ενώ στον πίνακα 2Γ περιγράφονται οι δομές των τριλέκτων που αναπτύσσονται εκατέρωθεν του προσδιοριζόμενου συνθετικού.

Στην πρώτη, τρίτη και πέμπτη στήλη περιγράφεται η γραμματική κατηγορία του 1ου, του 2ου συστατικού και του 3ου συστατικού, ενώ στη δεύτερη και στην τέταρτη στήλη καταγράφονται το άρθρο, που μπορεί να μεσολαβεί ανάμεσα στο 1ο και στο 2ο συστατικό, και η πρόθεση ή/και το άρθρο που μεσολαβούν ανάμεσα στο 3ο και στο 4ο συστατικό. Στην πέμπτη στήλη δίνεται χαρακτηριστικό παράδειγμα του περιγραφόμενου τρόπου σύνθεσης, στην έκτη στήλη η συχνότητα με την οποία εμφανίστηκε στο υλικό μας η περιγραφόμενη δομή και στην 7 στήλη περιγράφονται η ισοδύναμη γραμματική κατηγορία και η ορολογική κατηγορία των άμεσων συνθετικών – Προσδιοριζόμενου και προσδιοριστικού – με τη σειρά που εμφανίζονται στον τριλεκτικό όρο.

Πίνακας 2Α – Τρόποι σύνθεσης τριλεκτικών όρων (τριλέκτων) βιοϊατρικής με ανάπτυξη δεξιά του προσδιοριζόμενου συνθετικού

1	2	3	4	5	6	7	8
1° συστ.	Άρθρο	2° συστ.	Άρθρο/ Πρόθ.	3° συστ.	Παράδειγμα	Συχν.	Ορολ. Κατηγ.
Ο		Ε		Ο2	σύστημα ελεύθερο κυττάρων	3	ΟΥΣ_Ππ
Ο		Ε		ΞΑΚΡ	μεθυλάσες τροποποιητικές DNA	1	ΟΥΣ_Ππ
Ο		Ε	ΑΡΘ	Ο2	παράγοντες ευαισθητοποιοί της ακτινοβολίας	1	ΟΥΣ_Ππ
Ο		Ε	με	Ο3	πρωτεΐνες σχετικές με μικροσωληνάρια	1	ΟΥΣ_Ππ
Ο		MTX		Ο3	πρωτεΐνες δεσμεύουσες καλμοδουλίνη	1	ΟΥΣ_Ππ
Ο		MTX	με	Ο3	πρωτεΐνες συνδεδεμένες με ασβέστιο	1	ΟΥΣ_Ππ
Ο		MTX	με	ΞΑΚΡ	λέμφωμα σχετιζόμενο με AIDS	1	ΟΥΣ_Ππ
Ο		Ε2		Ο2	ιός δάγγειου πυρετού	232	ΟΥΣ_Ππ
Ο	ΑΡΘ	Ε2		Ο2	μετάθεση των	1	ΟΥΣ_Ππ

Ε.Π. «Κ.τ.Π.» Μέτρο 3.3 «Έρευνα και Τεχνολογική Ανάπτυξη στην ΚτΠ» Πράξη «Επεξεργασία εικόνων, ήχου και γλώσσας»	Έργο 9: «ΙΑΤΡΟΛΕΞΗ»
--	------------------------

					μεγάλων αγγείων		
Ο		Ε2		ΞΑΚΡ	υποδοχείς κυκλικού AMP	1	ΟΥΣ_Ππ
Ο		Ο2-		Ο2	εμβόλια τύφου-παράτυφου	11	ΟΥΣ_Ππ
Ο		Ο2-		ΞΑΚΡ	ρεδοκτάση φερεδοξίνης-NADP	1	ΟΥΣ_Ππ
Ο		ΞΛ-		ΞΛ	σύνδρομο Ehlers-Danlos	6	ΟΥΣ_Ππ
Ο		Ο2		Ο2	διαταραχές αντίληψης ήχου	1	ΟΥΣ_Ππ
Ο		Ο2	ΑΡΘ	Ο2	διαταραχές διαφοροποίησης του φύλου	4	ΟΥΣ_Ππ
Ο	ΑΡΘ	Ο2	ΑΡΘ	Ο2	νεύρο της χορδής του τυμπάνου	2	ΟΥΣ_Ππ
Ο		ΞΛ		Ο2	ιός Distemper σκύλου	2	ΟΥΣ_Ππ
Ο		ΞΛ		ΞΛ	σύμπλεγμα Mycobacterium Avium	1	ΟΥΣ_Ππ
Ο		Ο2		ΞΛ	ιός δάσους Semliki	6	ΟΥΣ_Ππ
Ο		Ο2		ΑΣ	υποδοχείς συμπληρώματος 3B	2	ΟΥΣ_Ππ
Ο		ΞΛ		Α	ιός Echo 9	1	ΟΥΣ_Ππ
Ο		Ο2		Α	γονίδια νευρινωμάτωσης 1	6	ΟΥΣ_Ππ
Ο		Σ		Ο2	ανοσοσφαιρίνες J αλυσίδας	3	ΟΥΣ_Ππ

Πίνακας 2B – Τρόποι σύνθεσης τριλεκτικών όρων (τριλέκτων) βιοϊατρικής με ανάπτυξη αριστερά του προσδιοριζόμενου συνθετικού

1	2	3	4	5	6
1° συστ.	2° συστ.	3° συστ.	Παράδειγμα	Συχν.	Ορολ. Κατηγ.
ΕΠΡ	MTX	Ο	βίαια εκπνεόμενος όγκος	4	ΟΥΣ_πΠ
Ε	Ε	Ο	άτυπες βακτηριακές μορφές	157	ΟΥΣ_πΠ
Ε-	Ε	Ο	ορθολογική-συναισθηματική ψυχοθεραπεία	1	ΟΥΣ_πΠ
Σ	Ε	Ο	γάμα κινητικοί νευρώνες	2	ΟΥΣ_πΠ
Σ	Ε	Ο	Τ ρυθμιστικά λεμφοκύτταρα	18	ΟΥΣ_πΠ
Σ-Α	Ε	Ο	ωμέγα-3 λιπαρά οξέα	1	ΟΥΣ_πΠ
Σ	Α	Ο	άλφα 1-αντιτρυψίνη	2	ΟΥΣ_πΠ
Ε	Α-Ε	Ο	μονοφωσφορική 8-βρωμοκυκλική αδενοσίνη	1	ΟΥΣ_πΠ

Ε.Π. «Κ.τ.Π.» Μέτρο 3.3 «Έρευνα και Τεχνολογική Ανάπτυξη στην ΚτΠ» Πράξη «Επεξεργασία εικόνων, ήχου και γλώσσας»	Έργο 9: «ΙΑΤΡΟΛΕΞΗ»
---	------------------------

E	A-E	O	κυτιδινομονοφωσφορικό N-ακετυλονευραμινικό οξύ	1	OYΣ_πΠ
E	Σ	O	αμυλοειδής βήτα πρωτεΐνη	1	OYΣ_πΠ
ΞΛ	E	O	Gram αρνητικά βακτήρια	3	OYΣ_πΠ
MOP	E	O	μη χρωμιόφιλα παραγάγγλια	9	OYΣ_πΠ
E	MOP	O	γενετική μη διάζευξη	1	OYΣ_πΠ

Πίνακας 2Γ – Τρόποι σύνθεσης τριλεκτικών όρων (τριλέκτων) βιοϊατρικής με ανάπτυξη εκατέρωθεν του προσδιοριζόμενου συνθετικού

1	2	3	4	5	6	7
1 ^ο συστ.	2 ^ο συστ.	Άρθρο	3 ^ο συστ.	Παράδειγμα	Συχν.	Ορολ. Κατηγ.
E	O		O2	συγκυτιακό ιοί αναπνευστικού	152	OYΣ_πΠ
E	O	ΑΡΘ	O2	πλασματοκυτταρικό κοκκίωμα του κόγχου	2	OYΣ_πΠ
E	O	ΑΡΘ	ΞΑΚΡ	περιοριστικά ένζυμα του DNA	1	OYΣ_πΠ
E	O		ΞΑΚΡ	ετερόγενο πυρηνικό RNA	2	OYΣ_πΠ
E	O		A	αιμοπεταλιακός παράγοντας 3	3	OYΣ_πΠ
ΣΥΝΟΛΟ					650	

Οι πίνακες 3Α, 3Β και 3Γ καλύπτουν όλες σχεδόν τις περιπτώσεις των τετραλεκτικών όρων που καταγράφηκαν στο υλικό μας. Πιο συγκεκριμένα, στον πίνακα 3Α δίνονται οι τρόποι σύνθεσης των τετραλέκτων με δομές που αναπτύσσονται δεξιά του προσδιοριζόμενου συνθετικού, που είναι μια ονοματική φράση. Στον πίνακα 3Β δίνονται οι τρόποι σύνθεσης των τετραλέκτων με δομές που αναπτύσσονται αριστερά του προσδιοριζόμενου συνθετικού, που είναι μια ονοματική φράση, ενώ στον πίνακα 3Γ περιγράφονται οι δομές των τετραλέκτων που αναπτύσσονται εκατέρωθεν του προσδιοριζόμενου συνθετικού.

Στην πρώτη, δεύτερη, τρίτη και τέταρτη ή πέμπτη² στήλη (ανάλογα με τον πίνακα) περιγράφεται η γραμματική κατηγορία του 1ου, του 2ου, του 3ου και του 4ου συστατικού. Στην πέμπτη ή έκτη στήλη (ανάλογα με τον πίνακα) δίνεται χαρακτηριστικό παράδειγμα του περιγραφόμενου τρόπου σύνθεσης, στην έκτη ή έβδομη στήλη η συχνότητα με την οποία εμφανίστηκε στο υλικό μας η περιγραφόμενη δομή και στην έβδομη ή στην όγδοη στήλη περιγράφονται η ισοδύναμη γραμματική κατηγορία και η ορολογική κατηγορία των άμεσων συνθετικών – Προσδιοριζόμενου και προσδιοριστικού – με τη σειρά που εμφανίζονται στον τετραλεκτικό όρο.

² Εξαίρεση γίνεται στον πίνακα 3Α, αφού στην τέταρτη στήλη περιγράφεται ο σύνδεσμος «και» στις δομές όπου καταγράφεται παράταξη του 3^{ου} και του 4^{ου} συστατικού.

Ε.Π. «Κ.Τ.Π.» Μέτρο 3.3 «Έρευνα και Τεχνολογική Ανάπτυξη στην ΚτΠ» Πράξη «Επεξεργασία εικόνων, ήχου και γλώσσας»	Έργο 9: «ΙΑΤΡΟΛΕΞΗ»
--	------------------------

Πίνακας 3Α – Τρόποι σύνθεσης τετραλεκτικών όρων (τετραλέκτων) βιοϊατρικής με ανάπτυξη δεξιά του προσδιοριζόμενου συνθετικού

1	2	3	4	5	6	7	8
1° συστ.	2° συστ.	3° συστ.		4° συστ.	Παράδειγμα	Συχν.	Ορολ. Κατηγ.
Ο	Ο2	Σ		Ο2	ιός ηπατίτιδας Β νήσας	1	Ουσ_Ππ
Ο	Ο2	Ε2		Ο2	απόφραξη αυχένα ουροδόχου κύστης	14	Ουσ_Ππ
Ο	Ε2	Ο2		Ο2	υλικά εσωτερικής επένδυσης οδοντοστοιχίας	10	Ουσ_Ππ
Ο	Ε2	Ε2		Ο2	γάγγλια αυτόνομου νευρικού συστήματος	14	Ουσ_Ππ
Ο	Α	Ε2		Ο2	αντιγόνο 1 λεμφοκυτταρικής λειτουργίας	1	Ουσ_Ππ
Ο	Ο2	Ο2-		Ο2	σφράγισμα οξειδίου ψευδαργύρου-ευγενόλης	3	Ουσ_Ππ
Ο	Ο2	Ο2	και	Ο2	υλικά σφράγισης ρωγμών και οπών	1	Ουσ_Ππ
Ο	ΕΠΡ	Ε2		Ο2	υποδοχείς πολύ όψιμου αντιγόνου	1	Ουσ_Ππ
Ο	ΜΤΧ	Ε2		Ο2	βοθρία επικαλυμμένα κυτταρικής μεμβράνης	1	Ουσ_Ππ

Πίνακας 3Β – Τρόποι σύνθεσης τετραλεκτικών όρων (τετραλέκτων) βιοϊατρικής με ανάπτυξη αριστερά του προσδιοριζόμενου συνθετικού

1	2	3	4	5	6	7
1° συστ.	2° συστ.	3° συστ.	4° συστ.	Παράδειγμα	Συχν.	Ορολ. Κατηγ.
ΕΠΡ	ΜΤΧ	Ε	Ο	φαρμακευτικά εξοπλισμένες ενδομητρικές συσκευές	1	Ουσ_πΠ
ΜΟΡ	Ε	Ε	Ο	μη στεροειδείς εκτρωτικοί παράγοντες	3	Ουσ_πΠ
Ε	ΜΟΡ	Ε	Ο	οξεία μη λεμφοκυτταρική λευχαιμία	1	Ουσ_πΠ

<p>Ε.Π. «Κ.τ.Π.» Μέτρο 3.3 «Έρευνα και Τεχνολογική Ανάπτυξη στην ΚτΠ» Πράξη «Επεξεργασία εικόνων, ήχου και γλώσσας»</p>	<p>Έργο 9: «ΙΑΤΡΟΛΕΞΗ»</p>
---	---------------------------------

Πίνακας 3Γ – Τρόποι σύνθεσης τετραλεκτικών όρων (τετραλέκτων) βιοϊατρικής με ανάπτυξη εκατέρωθεν του προσδιοριζόμενου συνθετικού

1	2	3	4	5	6	7
1° συστ.	2° συστ.	3° συστ.	4° συστ.	Παράδειγμα	Συχν.	Ορολ. Κατηγ.
Ε	Ο	Ε2	Ο2	ερευνητική εφαρμογή νέου φαρμάκου	30	Ουσ_Ππ
ΕΠΡ	ΜΤΧ	Ο	Ο2	άπω εσπειραμένα σωληνάρια νεφρού	2	Ουσ_πΠ
ΣΥΝΟΛΟ					83	

5. ΚΑΝΟΝΕΣ ΑΝΑΓΝΩΡΙΣΗΣ ΠΟΛΥΛΕΚΤΙΚΩΝ ΣΥΜΠΛΟΚΩΝ ΟΡΩΝ ΒΙΟΪΑΤΡΙΚΗΣ

Η λεπτομερής καταγραφή και ανάλυση των περίπου 213 διαφορετικών δομών των διλέκτων, τριλέκτων και τετραλέκτων του πεδίου της βιοϊατρικής επέτρεψε την εξαγωγή των περίπου είκοσι βασικών κανόνων σχηματισμού των πολυλεκτικών σύμπλοκων όρων του πεδίου, σύμφωνα με τους οποίους είναι δυνατή η αναζήτηση, ο εντοπισμός και η εξαγωγή τους από τα κείμενα.

Για την περιγραφή των κανόνων αναγνώρισης των πολυλεκτικών σύμπλοκων όρων της βιοϊατρικής χρησιμοποιήθηκε ο πιο διαδεδομένος φορμαλισμός περιγραφής της σύνταξης τεχνητών γλωσσών που είναι οι **definite-clause γραμματικές** (Pereira and Warren 1980). Στο πλαίσιο αυτού του φορμαλισμού ο κανόνας αναγνώρισης της ονοματικής φράσης (ΟΦ) μπορεί να περιγραφεί ως εξής:

ΟΦ(Γένος, Αριθμός, Πτώση) ←
 Άρθ(Γένος, Αριθμός, Πτώση) Ουσ(Γένος, Αριθμός, Πτώση)

Σύμφωνα με τον κανόνα αυτό ένα ουσιαστικό μετά από ένα άρθρο συνιστούν μία ονοματική φράση όταν συμφωνούν κατά γένος, αριθμό και πτώση. Επιπλέον, η ονοματική φράση έχει το ίδιο Γένος, Αριθμό και Πτώση με το άρθρο και το ουσιαστικό.

5.1 ΓΕΝΙΚΟΙ ΚΑΝΟΝΕΣ ΣΧΗΜΑΤΙΣΜΟΥ ΟΝΟΜΑΤΙΚΩΝ ΦΡΑΣΕΩΝ (ΟΦ)

Ακολουθεί ένα δείγμα στοιχειωδών definite-clause κανόνων που περιγράφουν το σχηματισμό των ονοματικών φράσεων που απαντούν σε κείμενα βιοϊατρικής. Ο κανόνας:

ΟΥΣ(Γένος, Αριθμός, Πτώση)←
 Ουσ(Γένος, Αριθμός, Πτώση) | Ξλξ | Ξακ | Ξαρ | Σμβ | Νμρ

ορίζει ότι σε θέση ουσιαστικού, εκτός από ένα κλιτό ουσιαστικό που φέρει τα μορφολογικά χαρακτηριστικά Γένος, Αριθμό και Πτώση, μπορεί επίσης να εμφανίζεται μία ξένη λέξη (Ξλ)³ ή ένα ξένο ακρώνυμο (Ξακ) ή ένα ξένο αρκτικόλεξο (Ξαρ) ή ένα σύμβολο (Σμβ) ή ένα νούμερο (Νμρ). Πιο συγκεκριμένα, το οροστοιχείο «Νούμερο» ορίζεται με τον παρακάτω definite-clause κανόνα:

³ Οι ξένες λέξεις μπορεί να είναι γραμμένες με ελληνικούς ή λατινικούς χαρακτήρες.

<p>Ε.Π. «Κ.τ.Π.» Μέτρο 3.3 «Έρευνα και Τεχνολογική Ανάπτυξη στην ΚτΠ» Πράξη «Επεξεργασία εικόνων, ήχου και γλώσσας»</p>	<p>Έργο 9: «ΙΑΤΡΟΛΕΞΗ»</p>
--	---------------------------------

Νμρ←
«1» | «2» | «3» |...

Αντίστοιχα, ο κανόνας:

ΕΠΘ(Γένος, Αριθμός, Πτώση)←
[Επρ] (Επθ(Γένος, Αριθμός, Πτώση) | Μπχ(Γένος, Αριθμός, Πτώση)) | Νμρ-
Επθ(Γένος, Αριθμός, Πτώση)

ορίζει ότι σε θέση επιθετικού προσδιορισμού μπορεί να εμφανίζεται α) ένα επίθετο ή μία μετοχή, με τα μορφολογικά χαρακτηριστικά Γένος, Αριθμό και Πτώση, που μπορεί να προσδιορίζεται από ένα επίρρημα με προαιρετική εμφάνιση που προηγείται, ή β) ένα επίθετο που είναι σύνθετο με ένα νούμερο (Νμρ) σε θέση προθήματος.

Οι κανόνες:

Αρθ_Οουσ(Γένος, Αριθμός, Πτώση)←
[Αρθ(Γένος, Αριθμός, Πτώση)] ΟΥΣ(Γένος, Αριθμός, Πτώση)

Αρθ_Επθ(Γένος, Αριθμός, Πτώση)←
[Αρθ(Γένος, Αριθμός, Πτώση)] ΕΠΘ(Γένος, Αριθμός, Πτώση)

δηλώνουν την προαιρετική παρουσία του άρθρου μπροστά από μία ονοματική φράση, που αποτελείται αντίστοιχα από ένα ουσιαστικό ή από ένα επίθετο.

5.2 ΓΕΝΙΚΟΙ ΚΑΝΟΝΕΣ ΣΧΗΜΑΤΙΣΜΟΥ ΠΡΟΘΕΤΙΚΩΝ ΦΡΑΣΕΩΝ (ΠΦ)

Οι προθετικές φράσεις περιγράφονται γενικώς από τον παρακάτω κανόνα:

ΠΦ(Γένος, Αριθμός, Πτώση)←
Πρθ ΟΦ (Γένος, Αριθμός, Πτώση=Γεν | Αιτ)

Σύμφωνα με τον κανόνα αυτό, μία προθετική φράση αποτελείται από μία πρόθεση και μία ονοματική φράση είτε σε γενική είτε σε αιτιατική (ανάλογα με την πρόθεση). Πιο συγκεκριμένα, για τον τομέα της βιοϊατρικής, το οροστοιχείο Πρόθεση ορίζεται με τον παρακάτω definite-clause κανόνα:

Πρθ←
«με» | «σε» | «από» | «κατά» | «για» | «μέσω» | «μεταξύ» | «έναντι»|...

5.3 ΚΑΝΟΝΕΣ ΑΝΑΓΝΩΡΙΣΗΣ ΠΟΛΥΛΕΚΤΙΚΩΝ ΣΥΜΠΛΟΚΩΝ ΟΡΩΝ ΒΙΟΪΑΤΡΙΚΗΣ

5.3.1 Διλεκτικοί όροι

Οι διλεκτικοί όροι αντιστοιχούν σε απλή ΟΦ (ΑΟΦ) ή σύνθετη ΟΦ (ΣΟΦ). Οι όροι που εντοπίστηκαν περιγράφονται με τους ακόλουθους κανόνες:

Κανόνας 1:

Άρθρ_Επθ_Ουσ(Γένος, Αριθμός, Πτώση)←

Άρθρ_Επθ(Γένος, Αριθμός, Πτώση) Ουσ(Γένος, Αριθμός, Πτώση)

Σύμφωνα με αυτό τον κανόνα ένας διλεκτικός όρος μπορεί να συντίθεται από ένα επίθετο και ένα ουσιαστικό με προαιρετική την παρουσία του άρθρου, π.χ. *(ο) ογκογόνος ιός, (το) βακτηριακό εμβόλιο, (το) τοπικό αναισθητικό.*

Κανόνας 2:

Ουσ_Ουσ(Γένος1, Αριθμός1, Πτώση1)←

Άρθρ_Ουσ(Γένος1, Αριθμός1, Πτώση1) Άρθρ_Ουσ ((Γένος2, Αριθμός2, Πτώση=Γεν) | (Γένος2, Αριθμός=Ενι, Πτώση=Γεν) | (Γένος2, Αριθμός=Πλη, Πτώση=Γεν))

Σύμφωνα με αυτό τον κανόνα ένας διλεκτικός όρος μπορεί να συντίθεται από δύο ονοματικές φράσεις εκ των οποίων η πρώτη είναι η κεφαλή και η δεύτερη είναι ετερόπτωτο προσδιοριστικό σε γενική της πρώτης, π.χ *(ο) ιός (της) λύσσας, (η) καλύπτρα του μεσεγκεφάλου, (οι) δίαυλοι ιόντων.*

Κανόνας 3:

ΟΟΦ(Γένος1, Αριθμός1, Πτώση1)←

ΟΦ(Γένος1, Αριθμός1, Πτώση1) ΟΥΣ ((Γένος2, Αριθμός2, Πτώση1) | (Γένος2, Αριθμός1, Πτώση1))

Σύμφωνα με τον παραπάνω κανόνα ένας διλεκτικός όρος μπορεί να συντίθεται από δύο ονοματικές φράσεις εκ των οποίων η πρώτη είναι η κεφαλή και η δεύτερη είναι ομοιόπτωτο προσδιοριστικό της, π.χ *(ο) βάκιλλος μεγαθήριος, (το) μέλος φάντασμα*

Κανόνας 4:

Ουσ_ΠΦ(Γένος1, Αριθμός1, Πτώση1)←

Άρθρ_Ουσ(Γένος1, Αριθμός1, Πτώση1) ΠΦ(Γένος2, Αριθμός2, Πτώση2)

Σύμφωνα με τον παραπάνω κανόνα ένας διλεκτικός όρος μπορεί να συντίθεται από μία ονοματική φράση και μία προθετική φράση, η οποία είναι προσδιοριστικό της πρώτης, π.χ *(η) αγγειοπλαστική με laser, (η) ευπάθεια σε νόσο.*

5.3.2 Τριλεκτικοί όροι

Οι τριλεκτικοί όροι αντιστοιχούν σε σύνθετη ΟΦ (ΣΟΦ). Παρατηρούνται ελάχιστοι όροι οι οποίοι είναι εκ γενετής τριλεκτικοί. Στις περισσότερες περιπτώσεις, η παραγωγή τους γίνεται με πυρήνα τους ήδη υπάρχοντες διλεκτικούς συνδυασμούς, οι οποίοι αναπτύσσονται με βάση τις εξής διαδικασίες:

1. Ανάπτυξη προς τα αριστερά:
 - α. με προσθήκη νέας κεφαλής: *δάγγειος πυρετός* → *ίος δάγγειου πυρετού, αντίληψη ήχου* → *διαταραχές αντίληψης ήχου*
 - β. με προσθήκη προσδιοριστή: *βακτηριακές μορφές* → *άτυπες βακτηριακές μορφές*
2. Συνένωση δύο διλεκτικών όρων: *διεγέρτης θυρεοειδούς + μακράς διάρκειας* → *διεγέρτης θυρεοειδούς μακράς διάρκειας*
3. Παράταξη δύο διλεκτικών όρων: *κισσοί οισοφάγου + κισσοί στομάχου* → *κισσοί οισοφάγου και στομάχου*

Με βάση τις προαναφερθείσες διαδικασίες και την τυπολογία επιτρεπτών δομών των διλεκτικών όρων, οι αντίστοιχες επιτρεπτές δομές για τους τριλεκτικούς όρους είναι αυτές που παρουσιάζονται στη συνέχεια:

Κανόνας 5:

Ουσ_Επθ_Ουσ(Γένος1, Αριθμός1, Πτώση1)←

Άρθρ_Ουσ(Γένος1, Αριθμός1, Πτώση1) Άρθρ_Επθ_Ουσ((Γένος2, Αριθμός2, Πτώση=Γεν) | (Γένος2, Αριθμός=Ενι, Πτώση=Γεν) | (Γένος2, Αριθμός=Πλη, Πτώση=Γεν))

Σύμφωνα με τον παραπάνω κανόνα ένας τριλεκτικός όρος μπορεί να αποτελείται από μία απλή ονοματική φράση (χωρίς προσδιοριστή) και μια επταυξημένη ονοματική φράση με επίθετο προσδιοριστή, η οποία είναι προσδιοριστικό της πρώτης, π.χ. *(το) μυοχαλαρωτικό κεντρικής δράσης, (το) εμβόλιο ιογενούς ηπατίτιδας, (η) μετάθεση των μεγάλων αγγείων.*

Κανόνας 6:

Επθ_Ουσ_Ουσ(Γένος1, Αριθμός1, Πτώση1) ←

Άρθρ_Επθ_Ουσ(Γένος1, Αριθμός1, Πτώση1) Άρθρ_Ουσ((Γένος2, Αριθμός2, Πτώση=Γεν) | (Γένος2, Αριθμός=Ενι, Πτώση=Γεν) | (Γένος2, Αριθμός=Πλη, Πτώση=Γεν))

Σύμφωνα με τον παραπάνω κανόνα ένας τριλεκτικός όρος μπορεί να αποτελείται από μία επταυξημένη με επίθετο προσδιοριστή ονοματική φράση και μια απλή

ονοματική φράση χωρίς προσδιοριστή, η οποία είναι προσδιοριστικό της πρώτης, π.χ. *(οι) υδρόφιλοι φακοί επαφής, (οι) πρώιμες πρωτεΐνες (των) αδενοϊών.*

Κανόνας 7:

Ουσ_Ουσ_Ουσ(Γένος1, Αριθμός1, Πτώση1)←

Άρθρ_Ουσ(Γένος1, Αριθμός1, Πτώση1) Άρθρ_Ουσ ((Γένος2, Αριθμός2, Πτώση=Γεν) | (Γένος2, Αριθμός=Ενι, Πτώση=Γεν) | (Γένος2, Αριθμός=Πλη, Πτώση=Γεν) Άρθρ_Ουσ ((Γένος3, Αριθμός2, Πτώση=Γεν) | (Γένος3, Αριθμός=Ενι, Πτώση=Γεν) | (Γένος3, Αριθμός=Πλη, Πτώση=Γεν))

Σύμφωνα με τον παραπάνω κανόνα ένας τριλεκτικός όρος μπορεί να αποτελείται από μία απλή ονοματική φράση, που προσδιορίζεται από ένα διλεκτικό της μορφής Κανόνα 2, π.χ. *(οι) διαταραχές αντίληψης ήχου, (ο) παράγοντας πήξης αίματος.*

Κανόνας 8:

Επθ_Επθ_Ουσ(Γένος, Αριθμός, Πτώση) ←

Άρθρ_Επθ(Γένος, Αριθμός, Πτώση) ΕΠΘ(Γένος, Αριθμός, Πτώση) ΟΥΣ(Γένος, Αριθμός, Πτώση)

Σύμφωνα με τον παραπάνω κανόνα ένας τριλεκτικός όρος μπορεί να αποτελείται από ένα διλεκτικό της μορφής του Κανόνα 1, που προσδιορίζεται από ένα επιπλέον επίθετο, π.χ. *(το) κεντρικό νευρικό σύστημα, (η) άτυπη βακτηριακή μορφή.*

Κανόνας 9:

Ουσ_Ουσ_Συνδ_Ουσ(Γένος1, Αριθμός1, Πτώση1)←

Άρθρ_Ουσ(Γένος1, Αριθμός1, Πτώση1) Άρθρ_Ουσ ((Γένος2, Αριθμός2, Πτώση=Γεν) | (Γένος2, Αριθμός=Ενι, Πτώση=Γεν) | (Γένος2, Αριθμός=Πλη, Πτώση=Γεν) Συνδ Άρθρ_Ουσ ((Γένος3, Αριθμός2, Πτώση=Γεν) | (Γένος3, Αριθμός=Ενι, Πτώση=Γεν) | (Γένος3, Αριθμός=Πλη, Πτώση=Γεν))

Σύμφωνα με τον παραπάνω κανόνα ένας τριλεκτικός όρος σχηματίζεται με παράταξη δύο διλεκτικών της μορφής του Κανόνα 2, που διατηρούν την ίδια κεφαλή, π.χ. *κισσοί οισοφάγου και στομάχου.*

Κανόνας 10:

Επθ_Ουσ_ΠΦ(Γένος1, Αριθμός1, Πτώση1) ←

Άρθρ_Επθ_Ουσ(Γένος1, Αριθμός1, Πτώση1) ΠΦ(Γένος2, Αριθμός2, Πτώση2)

Σύμφωνα με τον παραπάνω κανόνα ένας τριλεκτικός όρος σχηματίζεται με σύνθεση ενός διλεκτικού όρου της μορφής του Κανόνα 1 και μιας προθετικής φράσης, π.χ. *(οι) μετατραυματικές διαταραχές από στρες, (η) τροφική δηλητηρίαση από σταφυλόκοκκο.*

Κανόνας 11:

Ουσ_Ουσ_ΠΦ(Γένος1, Αριθμός1, Πτώση1)←

Άρθρ_Ουσ(Γένος1, Αριθμός1, Πτώση1) Άρθρ_Ουσ ((Γένος2, Αριθμός2, Πτώση=Γεν) | (Γένος2, Αριθμός=Ενι, Πτώση=Γεν) | (Γένος2, Αριθμός=Πλη, Πτώση=Γεν) ΠΦ(Γένος3, Αριθμός3, Πτώση3)

Σύμφωνα με τον παραπάνω κανόνα ένας τριλεκτικός όρος σχηματίζεται με σύνθεση ενός διλεκτικού όρου της μορφής του Κανόνα 2 και μιας προθετικής φράσης, π.χ. *(οι) ιοί εγκεφαλίτιδας από κρότωνες, (το) εμβόλιο πολιομυελίτιδας από το στόμα.*

5.3.3 Τετραλεκτικοί όροι

Η παραγωγή των τετραλεκτικών όρων γίνεται με πυρήνα τους ήδη υπάρχοντες διλεκτικούς και τριλεκτικούς συνδυασμούς, οι οποίοι αναπτύσσονται με βάση τις εξής διαδικασίες:

1. Ανάπτυξη προς τα αριστερά:
 - α. με προσθήκη νέας κεφαλής: *αυτόνομο νευρικό σύστημα → γάγγλια αυτόνομου νευρικού συστήματος*
 - β. με προσθήκη προσδιοριστή: *εφαρμογή νέου φαρμάκου → ερευνητική εφαρμογή νέου φαρμάκου*
2. Παράταξη δύο τριλεκτικών όρων: *λέμφωμα μικρών κυττάρων + λέμφωμα μεγάλων κυττάρων → λέμφωμα μικρών και μεγάλων κυττάρων*

Με βάση τις προαναφερθείσες διαδικασίες και την τυπολογία επιτρεπτών δομών των διλεκτικών και τριλεκτικών όρων, οι αντίστοιχες επιτρεπτές δομές για τους τετραλεκτικούς όρους είναι αυτές που παρουσιάζονται στη συνέχεια:

Κανόνας 12:

Ουσ_Ουσ_Επθ_Ουσ(Γένος1, Αριθμός1, Πτώση1)←

Άρθρ_Ουσ(Γένος1, Αριθμός1, Πτώση1) Άρθρ_Ουσ ((Γένος2, Αριθμός2, Πτώση=Γεν) | (Γένος2, Αριθμός=Ενι, Πτώση=Γεν) | (Γένος2, Αριθμός=Πλη, Πτώση=Γεν) Άρθρ_Επθ_Ουσ ((Γένος3, Αριθμός2, Πτώση=Γεν) | (Γένος3, Αριθμός=Ενι, Πτώση=Γεν) | (Γένος3, Αριθμός=Πλη, Πτώση=Γεν))

Σύμφωνα με τον παραπάνω κανόνα ένας τετραλεκτικός όρος μπορεί να αποτελείται από μία απλή ονοματική φράση (χωρίς προσδιοριστή) και μια τριλεκτική ονοματική φράση της μορφής Ουσ_Επθ_Ουσ (Κανόνας 5), η οποία είναι προσδιοριστικό της

πρώτης, π.χ (η) απόφραξη αυχένα ουροδόχου κύστης, (ο) χρόνος πήξης ολικού αίματος, (τα) υλικά πλήρωσης ριζικού σωλήνα.

Κανόνας 13:

Ουσ_Επθ_Ουσ_Ουσ(Γένος1, Αριθμός1, Πτώση1)←

Άρθρ_Ουσ(Γένος1, Αριθμός1, Πτώση1) Επθ_Επθ_Ουσ ((Γένος2, Αριθμός2, Πτώση=Γεν) | (Γένος2, Αριθμός=Ενι, Πτώση=Γεν) | (Γένος2, Αριθμός=Πλη, Πτώση=Γεν)

Σύμφωνα με τον παραπάνω κανόνα ένας τετραλεκτικός όρος μπορεί να αποτελείται από μία απλή ονοματική φράση (χωρίς προσδιοριστή) και μια τριλεκτική ονοματική φράση της μορφής Επθ_Επθ_Ουσ (Κανόνας 8), η οποία είναι προσδιοριστικό της πρώτης, π.χ (τα) γάγγλια αυτόνομου νευρικού συστήματος, (τα) νεοπλάσματα άγνωστης πρωτοπαθούς εστίας.

Κανόνας 14:

Ουσ_Επθ_Ουσ_Ουσ(Γένος1, Αριθμός1, Πτώση1)←

Άρθρ_Ουσ(Γένος1, Αριθμός1, Πτώση1) Άρθρ_Επθ_Ουσ ((Γένος2, Αριθμός2, Πτώση=Γεν) | (Γένος2, Αριθμός=Ενι, Πτώση=Γεν) | (Γένος2, Αριθμός=Πλη, Πτώση=Γεν) Άρθρ_Ουσ ((Γένος3, Αριθμός2, Πτώση=Γεν) | (Γένος3, Αριθμός=Ενι, Πτώση=Γεν) | (Γένος3, Αριθμός=Πλη, Πτώση=Γεν))

Σύμφωνα με τον παραπάνω κανόνα ένας τετραλεκτικός όρος μπορεί να αποτελείται από μία απλή ονοματική φράση (χωρίς προσδιοριστή) και μια τριλεκτική ονοματική φράση της μορφής Επθ_Ουσ_Ουσ (Κανόνας 6), η οποία είναι προσδιοριστικό της πρώτης, π.χ. (τα) υλικά εσωτερικής επένδυσης οδοντοστοιχίας, (οι) πρωτεΐνες εξωτερικής μεμβράνης βακτηρίων.

Κανόνας 15:

Επθ_Ουσ_Επθ_Ουσ(Γένος1, Αριθμός1, Πτώση1) ←

Άρθρ_Επθ_Ουσ(Γένος1, Αριθμός1, Πτώση1) Άρθρ_Επθ_Ουσ((Γένος2, Αριθμός2, Πτώση=Γεν) | (Γένος2, Αριθμός=Ενι, Πτώση=Γεν) | (Γένος2, Αριθμός=Πλη, Πτώση=Γεν))

Σύμφωνα με τον παραπάνω κανόνα ένας τετραλεκτικός όρος μπορεί να αποτελείται από μία τριλεκτική ονοματική φράση της μορφής Ουσ_Επθ_Ουσ (Κανόνας 5), η οποία είναι επαυξημένη με ένα επίθετο προσδιοριστή, π.χ. (η) ερευνητική εφαρμογή νέου φαρμάκου, (οι) δραστικές ουσίες θειοβαρβιτουρικού οξέος.

Κανόνας 16:

Ουσ_Ουσ_Ουσ_ΠΦ(Γένος1, Αριθμός1, Πτώση1)←

Άρθρ_Ουσ(Γένος1, Αριθμός1, Πτώση1) Άρθρ_Ουσ ((Γένος2, Αριθμός2, Πτώση=Γεν) | (Γένος2, Αριθμός=Ενι, Πτώση=Γεν) | (Γένος2, Αριθμός=Πλη, Πτώση=Γεν) Άρθρ_Ουσ ((Γένος3, Αριθμός3, Πτώση=Γεν) | (Γένος3, Αριθμός=Ενι, Πτώση=Γεν) | (Γένος3, Αριθμός=Πλη, Πτώση=Γεν) ΠΦ(Γένος4, Αριθμός4, Πτώση4)

Σύμφωνα με τον παραπάνω κανόνα ένας τετραλεκτικός όρος μπορεί να αποτελείται από μία τριλεκτική ονοματική φράση της μορφής Ουσ_Ουσ_Ουσ (Κανόνας 7), η οποία είναι επαυξημένη κατά μία προθετική φράση, π.χ. *(ο) ιός πυρετού Κολοράντο από κρότωνες.*

Κανόνας 17:

Ουσ_Επθ_Ουσ_ΠΦ(Γένος1, Αριθμός1, Πτώση1)←

Άρθρ_Ουσ(Γένος1, Αριθμός1, Πτώση1) Άρθρ_Επθ_Ουσ ((Γένος2, Αριθμός2, Πτώση=Γεν) | (Γένος2, Αριθμός=Ενι, Πτώση=Γεν) | (Γένος2, Αριθμός=Πλη, Πτώση=Γεν) ΠΦ(Γένος3, Αριθμός3, Πτώση3)

Σύμφωνα με τον παραπάνω κανόνα ένας τετραλεκτικός όρος μπορεί να αποτελείται από μία τριλεκτική ονοματική φράση της μορφής Ουσ_Επθ_Ουσ (Κανόνας 5), η οποία είναι επαυξημένη κατά μία προθετική φράση, π.χ. *(η) διαταραχή μειωμένης προσοχής με υπερκινητικότητα.*

Κανόνας 18:

Επθ_Επθ_Ουσ_ΠΦ(Γένος, Αριθμός, Πτώση) ←

Άρθρ_Επθ(Γένος, Αριθμός, Πτώση) ΕΠΘ(Γένος, Αριθμός, Πτώση) ΟΥΣ(Γένος, Αριθμός, Πτώση) ΠΦ(Γένος2, Αριθμός2, Πτώση2)

Σύμφωνα με τον παραπάνω κανόνα ένας τετραλεκτικός όρος μπορεί να αποτελείται από μία τριλεκτική ονοματική φράση της μορφής Επθ_Επθ_Ουσ (Κανόνας 8), η οποία είναι επαυξημένη κατά μία προθετική φράση.

Κανόνας 19:

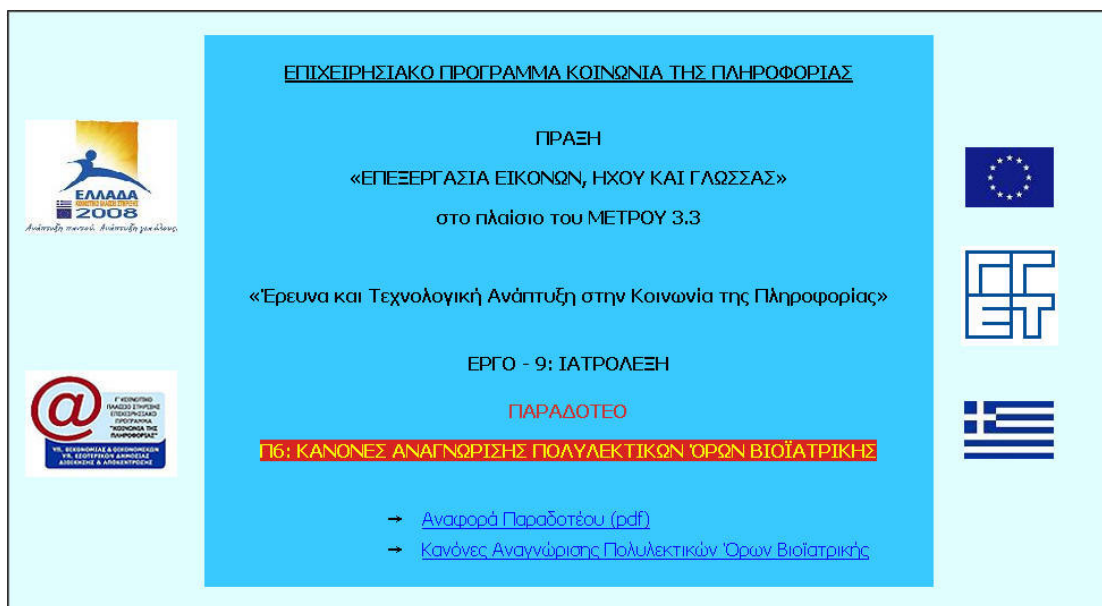
Επθ_Ουσ_Ουσ_ΠΦ(Γένος1, Αριθμός1, Πτώση1) ←

Άρθρ_Επθ_Ουσ(Γένος1, Αριθμός1, Πτώση1) Άρθρ_Ουσ((Γένος2, Αριθμός2, Πτώση=Γεν) | (Γένος2, Αριθμός=Ενι, Πτώση=Γεν) | (Γένος2, Αριθμός=Πλη, Πτώση=Γεν)) ΠΦ(Γένος3, Αριθμός3, Πτώση3)

Σύμφωνα με τον παραπάνω κανόνα ένας τετραλεκτικός όρος μπορεί να αποτελείται από μία τριλεκτική ονοματική φράση της μορφής Επθ_Ουσ_Ουσ (Κανόνας 6), η οποία είναι επαυξημένη κατά μία προθετική φράση.

6. ΟΔΗΓΙΕΣ ΧΡΗΣΗΣ ΤΟΥ ΟΠΤΙΚΟΥ ΔΙΣΚΟΥ

Στο παρόν παραδοτέο περιλαμβάνεται οπτικός δίσκος (CD) ο οποίος περιέχει τους μονολεκτικούς όρους βιοϊατρικής, καθώς και την παρούσα τεκμηρίωση. Μόλις εισαχθεί ο οπτικός δίσκος στη μονάδα οπτικού δίσκου του υπολογιστή θα ανοίξει αυτόματα ο φυλλομετρητής με την παρουσίαση του παραδοτέου (εικόνα 1).



Εικόνα 1: Στιγμιότυπο παρουσίασης του παραδοτέου.

Στην πρώτη σελίδα παρουσιάζονται στοιχεία του έργου και του συγκεκριμένου παραδοτέου καθώς και δύο υπερσύνδεσμοι (links). Ο πρώτος οδηγεί στην αναφορά του παραδοτέου (παρούσα αναφορά) σε μορφή PDF και ο δεύτερος στους κανόνες αναγνώρισης Πολυλεκτικών όρων.

7. ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Adar, E., *S-RAD: A Simple and Robust Abbreviation Dictionary*. 2002, HP Lab.
2. Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, et al., *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs*. *Nucleic Acids Res*, 1997. **25**(17): p. 3389-402.
3. Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman, *Basic local alignment search tool*. *J Mol Biol*, 1990. **215**(3): p. 403-10.
4. Ananiadou, S. *A Methodology for Automatic Term Recognition*. In: *Proceedings of COLING-94*. 1994. Kyoto, Japan. p. 1034-1038.
5. Ananiadou, S., S. Albert, and D. Schuhmann, *Evaluation of Automatic Term Recognition of Nuclear Receptors from Medline*. *Genome Informatics Series*, 2000.
6. Andrade, M.A. and A. Valencia, *Automatic Extraction of Keywords from Scientific Text: Application to the Knowledge Domain of Protein Families*. *Bioinformatics*, 1998. **14**(7): p. 600-7.
7. Aronson, A.R. *Effective mapping of biomedical text to the UMLS Metathesaurus: the MetaMap program*. In: *Proceedings of AMIA Symp*. 2001. p. 17-21.
8. Benson, D.A., I. Karsch-Mizrachi, D.J. Lipman, J. Ostell, B.A. Rapp, and D.L. Wheeler, *GenBank*. *Nucleic Acids Res*, 2000. **28**(1): p. 15-8.
9. Blake, C. and W. Pratt. *Better Rules, Fewer Features: A Semantic Approach to Selecting Features from Text*. In: *Proceedings of IEEE Data Mining Conference*. 2001. San Jose, California. p. 59-66.
10. Blaschke, C. and A. Valencia, *Molecular biology nomenclature thwarts information-extraction progress*. *IEEE Intelligent Systems*, 2002. **17**(3): p. 73-76. 74.
11. Boeckmann, B., A. Bairoch, R. Apweiler, M.C. Blatter, A. Estreicher, E. Gasteiger, et al., *The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003*. *Nucleic Acids Res*, 2003. **31**(1): p. 365-70.
12. Bourigault, D., I. Gomzalez-Mullier, and C. Gros. *LEXTER, a Natural Language Processing Tool for Terminology Extraction*. In: *Proceedings of EURALEX '96*. 1996. p. 771-9.
13. Brill, E. *A simple rule-based part-of-speech tagger*. In: *Proceedings of ANLP-92*. 1992. Trento, IT. p. 152-155.
14. Chang, J.T., H. Schutze, and R.B. Altman, *Creating an online dictionary of abbreviations from MEDLINE*. *J Am Med Inform Assoc*, 2002. **9**(6): p. 612-20.
15. Cohen, K.B., G.K. Acquah-Mensah, A.E. Dolbey, and L. Hunter. *Contrast and Variability in Gene Names*. In: *Proceedings of Workshop on NLP in the Biomedical Domain, ACL 2002*. 2002. Philadelphia. p. 14-20.
16. Collier, N., C. Nobata, and J. Tsujii. *Automatic Term Identification and Classification in Biological Texts*. In: *Proceedings of Natural Language Pacific Rim Symposium*. 1999. Beijing, China. p. 369-374.

- 17 Collier, N., C. Nobata, and J. Tsujii. *Extracting the Names of Genes and Gene Products with a Hidden Markov Model*. In: Proceedings of COLING 2000. 2000. Saarbruecken. p. 201-207.
- 18 Craven, M. and J. Kumlien. *Constructing biological knowledge bases by extracting information from text sources*. In: Proceedings of Int Conf Intell Syst Mol Biol. 1999. p. 77-86.
- 19 Frantzi, K., S. Ananiadou, and H. Mima, *Automatic Recognition of Multi-Word Terms: The Cvalue/NC-value method*. International Journal on Digital Libraries, 2000. **3**(2): p. 115-130.
- 20 Franzen, K., G. Eriksson, F. Olsson, L. Asker, P. Liden, and J. Coster, *Protein names and how to find them*. Int J Med Inf, 2002. **67**(1-3): p. 49-61.
- 21 Fukuda, K., A. Tamura, T. Tsunoda, and T. Takagi. *Toward information extraction: identifying protein names from biological papers*. In: Proceedings of Pac Symp Biocomput. 1998. p. 707-718.
- 22 Gaizauskas, R., G. Demetriou, and K. Humphreys. *Term Recognition and Classification in Biological Science Journal Articles*. In: Proceedings of Workshop on Computational Terminology for Medical and Biological Applications. 2000. Patras, Greece. p. 37-44.
- 23 Gaizauskas, R., G. Demetriou, P.J. Artymiuk, and P. Willett, *Protein structures and information extraction from biological texts: the PASTA system*. Bioinformatics, 2003. **19**(1): p. 135-43.
- 24 Harris, M.A., J. Clark, A. Ireland, J. Lomax, M. Ashburner, R. Foulger, et al., *The Gene Ontology (GO) database and informatics resource*. Nucleic Acids Res, 2004. **32**(1): p. D258-61.
- 25 Hatzivassiloglou, V., P.A. Duboue, and A. Rzhetsky, *Disambiguating Proteins, Genes, and RNA in Text: A Machine Language Approach*. Bioinformatics, 2001. **17 Suppl 1**: p. S97-106.
- 26 Hirschman, L., A.A. Morgan, and A.S. Yeh, *Rutabaga by any other name: extracting biological names*. J Biomed Inform, 2002. **35**(4): p. 247-59.
- 27 Hisamitsu, T. and J. Tsujii, *Measuring Term Representativeness*. Information Extraction in the Web Era, ed. M.T. Pazienza. 2003: LNAI 2700, Springer. 45-76.
- 28 Hobbs, J.R., D. Appelt, J. Bear, D. Israel, M. Kameyama, M. Stickel, et al., *FASTUS: A Cascaded Finite-State Transducer for Extracting Information From Natural-Language Text*. Finite-State Language Processing. 1997, Cambridge: MIT press. 383-406.
- 29 Hobbs, J.R., *Information extraction from biomedical text*. J Biomed Inform, 2002. **35**(4): p. 260-4.
- 30 Hodges, P.E., W.E. Payne, and J.I. Garrels, *The Yeast Protein Database (YPD): a curated proteome database for Saccharomyces cerevisiae*. Nucleic Acids Res, 1998. **26**(1): p. 68-72.
- 31 Hou, W. and H. Chen. *Enhancing Performance of Protein Name Recognizers Using Collocation*. In: Proceedings of NLP in Biomedicine, ACL 2003. 2003. Sapporo, Japan. p. 25-32.

- 32 Humphreys, B.L., D.A. Lindberg, H.M. Schoolman, and G.O. Barnett, *The Unified Medical Language System: an informatics research collaboration*. J Am Med Inform Assoc, 1998. **5**(1): p.1-11.
- 33 Humphreys, K., G. Demetriou, and R. Gaizauskas. *Two applications of information extraction to biological science journal articles: enzyme interactions and protein structures*. In: Proceedings of Pac Symp Biocomput. 2000. p. 505-516.
- 34 Jacquemin, C. and E. Tzoukermann, *NLP for Term Variant Extraction: A Synergy of Morphology, Lexicon and Syntax*, in *Natural Language Information Retrieval*, T. Strzalkowski, Editor. 1999, Kluwer: Boston. p. 25-74.
- 35 Jacquemin, C., *Spotting and Discovering Terms through NLP*. 2001, Cambridge, Ma: MIT Press.
- 36 Kazama, J., T. Makino, Y. Ohta, and J. Tsujii. *Tuning support vector machines for biomedical named entity recognition*. In: Proceedings of Workshop on NLP in the Biomedical Domain, ACL 2002. 2002. Philadelphia. p. 1-8.
- 37 ISO 1087-1:2000 *Terminology work -- Vocabulary -- Part 1: Theory and application*
- 38 Krauthammer, M., A. Rzhetsky, P. Morozov, and C. Friedman, *Using BLAST for identifying gene and protein names in journal articles*. Gene, 2000. **259**(1-2): p. 245-52.
- 39 Lee, K., Y. Hwang, and H. Rim. *Two-Phase Biomedical NE Recognition based on SVMs*. In: Proceedings of NLP in Biomedicine, ACL 2003. 2003. Sapporo, Japan. p. 33-40.
- 40 Liu, H., A.R. Aronson, and C. Friedman. *A study of abbreviations in MEDLINE abstracts*. In: Proceedings of AMIA Symp. 2002. p. 464-8.
- 41 Liu, H., S.B. Johnson, and C. Friedman, *Automatic resolution of ambiguous terms based on machine learning and conceptual relations in the UMLS*. J Am Med Inform Assoc, 2002. **9**(6): p.621-36.
- 42 Morgan, A., A. Yeh, L. Hirschman, and M. Colosimo. *Gene Name Extraction Using FlyBase resources*. In: Proceedings of NLP in Biomedicine, ACL 2003. 2003. Sapporo, Japan. p. 1-8.
- 43 Narayanaswamy, M., K.E. Ravikumar, and K. Vijay-Shanker. *A biological named entity recognizer*. In: Proceedings of Pac Symp Biocomput. 2003. p. 427-38.
- 44 Nenadic, G., I. Spasic, and S. Ananiadou, *Terminology-driven mining of biomedical literature*. Bioinformatics, 2003. **19**(8): p. 938-43.
- 45 Nenadic, G., I. Spasic, and S. Ananiadou. *Automatic Acronym Acquisition and Term Variation Management within Domain-Specific Texts*. In: Proceedings of LREC-3. 2002. Las Palmas, Spain. p. 2155-2162.
- 46 Nenadic, G., I. Spasic, and S. Ananiadou. *Mining Biomedical Abstracts: What is in a Term?* In: Proceedings of International Joint Conference on NLP. 2004. Sanya, China. p. 247-254.
- 47 Nenadic, G., S. Rice, I. Spasic, S. Ananiadou, and B.J. Stapley. *Selecting*

- Text Features for Gene Name Classification: from Documents to Terms.* In: *Proceedings of NLP in Biomedicine, ACL 2003.* 2003. Sapporo, Japan. p. 121-128.
- 48 Nobata, C., N. Collier, and J. Tsujii. *Automatic term identification and classification in biological texts.* In: *Proceedings of Natural Language Pacific Rim Symposium.* 1999. p. 369-374.
- 49 Ogren, P., K. Cohen, G. Acquaaah-Mensah, J. Eberlein, and L. Hunter. *The Compositional Structure of Gene Ontology Terms.* In: *Proceedings of Pac Symp Biocomput.* 2004. p. 214-25.
- 50 Ohta, T., Y. Tateisi, H. Mima, and J. Tsujii. *GENIA corpus: an Annotated Research Abstract Corpus in Molecular Biology Domain.* In: *Proceedings of Human Language Technology Conference (HLT 2002).* 2002. p. 73-7.
- 51 Pakhomov, S. *Semi-Supervised Maximum Entropy Based Approach to Acronym and Abbreviation Normalization in Medical Texts.* In: *Proceedings of 40th annual meeting of ACL.* 2002. p. 160-167.
- 52 Pereira, F. and Warren, D.H.D. *Definite Clause Grammars for Language Analysis – A survey of the formalism and a comparison with augmented transition networks,* *Journal of Artificial Intelligence,* 1980, vol. 13, no 3, pp. 231-278.
- 53 Proux, D., F. Rechenmann, L. Julliard, V.V. Pillet, and B. Jacq. *Detecting Gene Symbols and Names in Biological Texts: A First Step toward Pertinent Information Extraction.* In: *Proceedings of Ninth Workshop on Genome Informatics.* 1998. p. 72-80.
- 54 Pruitt, K.D. and D.R. Maglott, *RefSeq and LocusLink: NCBI gene-centered resources.* *Nucleic Acids Res,* 2001. **29**(1): p. 137-40.
- 55 Pustejovsky, J., J. Gastano, B. Cochran, M. Kotecki, M. Morrell, and A. Rumshisky. *Extraction and Disambiguation of Acronym-Meaning Pairs in Medline.* In: *Proceedings of Medinfo.* 2001.
- 56 Raychaudhuri, S., J.T. Chang, P.D. Sutphin, and R.B. Altman, *Associating genes with gene ontology codes using a maximum entropy analysis of biomedical literature.* *Genome Res,* 2002. **12**(1): p. 203-14.
- 57 Rimer, M. and M. O'Connell, *BioABACUS: a database of abbreviations and acronyms in biotechnology and computer science.* *Bioinformatics,* 1998. **14**(10): p. 888-9.
- 58 Rindfleisch, T.C., L. Hunter, and A.R. Aronson. *Mining molecular binding terminology from biomedical text.* In: *Proceedings of AMIA Symp.* 1999. p. 127-31.
- 59 Rindfleisch, T.C., L. Tanabe, J.N. Weinstein, and L. Hunter. *EDGAR: extraction of drugs, genes and relations from the biomedical literature.* In: *Proceedings of Pac Symp Biocomput.* 2000. p.517-28.
- 60 Schwartz, A.S. and M.A. Hearst. *A simple algorithm for identifying abbreviation definitions in biomedical text.* In: *Proceedings of Pac Symp Biocomput.* 2003. p. 451-62.
- 61 Seewald, A. *Towards Recognizing Domain and Species from MEDLINE Publications.* In: *Proceedings of European Workshops on Data Mining and*

- Text Mining for Bioinformatics*. 2003. p. 51-58.
- 62 Shen, D., J. Zhang, G. Zhou, J. Su, and C. Tan. *Effective Adaptation of Hidden Markov Modelbased Named Entity Recognizer for Biomedical Domain*. In: *Proceedings of NLP in Biomedicine, ACL 2003*. 2003. Sapporo, Japan. p. 49-56.
- 63 Spasic, I., G. Nenadic, and S. Ananiadou. *Using Domain-Specific Verbs for Term Classification*. In: *Proceedings of NLP in Biomedicine, ACL 2003*. 2003. Sapporo, Japan. p. 17-24.
- 64 Takeuchi, K. and N. Collier. *Bio-medical Entity Extraction using Support Vector Machines*. In: *Proceedings of NLP in Biomedicine, ACL 2003*. 2003. Sapporo, Japan. p. 57-64.
- 65 Tanabe, L. and W.J. Wilbur, *Tagging gene and protein names in biomedical text*. *Bioinformatics*, 2002. **18**(8): p. 1124-32.
- 66 Thomas, J., D. Milward, C. Ouzounis, S. Pulman, and M. Carroll. *Automatic extraction of protein interactions from scientific abstracts*. In: *Proceedings of Pac Symp Biocomput*. 2000. p. 541-552.
- 67 Torii, M. and K. Vijay-Shanker. *Using Unlabeled MEDLINE Abstracts for Biological Named Entity Classification*. In: *Proceedings of Genome Informatics Workshop 2002*. 2002. p. 567-658.
- 68 Torii, M., S. Kamboj, and K. Vijay-Shanker. *An Investigation of Various Information Sources for Classifying Biological Names*. In: *Proceedings of NLP in Biomedicine, ACL 2003*. 2003. Sapporo, Japan. p. 113-120.
- 69 Tsuruoka, Y. and J. Tsujii. *Boosting Precision and Recall of Dictionary-Based Protein Name Recognition*. In: *Proceedings of NLP in Biomedicine, ACL 2003*. 2003. Sapporo, Japan. p. 41-48.
- 70 Tsuruoka, Y. and J. Tsujii. *Probabilistic Term Variant Generator for Biomedical Terms*. In: *Proceedings of 26th Annual ACM SIGIR Conference*. 2003. p. 167-173.
- 71 Tuason, O., L. Chen, H. Liu, J.A. Blake, and C. Friedman. *Biological Nomenclature: A Source of Lexical Knowledge and Ambiguity*. In: *Proceedings of Pac Symp Biocomput*. 2004. p. 238-49.
- 72 Yamamoto, K., T. Kudo, A. Konagaya, and Y. Matsumoto. *Protein Name Tagging for Biomedical Annotation in Text*. In: *Proceedings of NLP in Biomedicine, ACL 2003*. 2003. Sapporo, Japan. p.65-72.
- 73 Yoshida, M., K. Fukuda, and T. Takagi, *PNAD-CSS: A Workbench for Constructing a Protein Name Abbreviation Dictionary*. *Bioinformatics*, 2000. **16**(2): p. 169-175.
- 74 Yu, H. and E. Agichtein, *Extracting Synonymous Gene and Protein Terms from Biological Literature*. *Bioinformatics*, 2003. **19 Suppl 1**: p. I340-I349.
- 75 Yu, H., G. Hripcsak, and C. Friedman, *Mapping abbreviations to full forms in biomedical articles*. *J Am Med Inform Assoc*, 2002. **9**(3): p. 262-72.
- 76 Ανανιάδου, Σ., Ζερβανού, Κ., *Αναγνώριση όρων σε υπολογιστικά συστήματα: προβλήματα και μέθοδοι*. Στο *Ελληνική ορολογία: έρευνα και εφαρμογές* (Κατσογιάννου, Μ., Ευθυμίου Ελένη, επιμ.), Αθήνα 2004, σ. 283-312.

- 77 Βαλεοντής Κ., Ζερίτη Κ., Νικολάκη Κ., *Το προσδιοριστικό συνθετικό του ελληνικού σύμπλοκου όρου και η χρήση της γενικής*, Πρακτικά 2ου Συνεδρίου «Ελληνική Γλώσσα και Ορολογία», Αθήνα 1999, σ. 283-316.
- 78 Γαβριηλίδου, Μ., Λαμπροπούλου, Π., *Εξαγωγή όρων από κείμενα: μια υβριδική μέθοδος*. Στο *Ελληνική ορολογία: έρευνα και εφαρμογές* (Κατσογιάννου, Μ., Ευθυμίου Ελένη, επιμ.), Αθήνα 2004, σ. 313-326.