

ΙΑΤΡΟΛΕΞΗ

Το ορολόγιο της βιοϊατρικής: συλλογή και επεξεργασία του γλωσσικού υλικού

Μαβίνα Πανταζάρα, Δρ. Γλωσσολόγος

Η μακροδομή της οντολογίας

Ένα από τα βασικά βήματα για την ανάπτυξη της οντολογίας των βιοϊατρικών όρων αποτελεί η κατάρτιση της **μακροδομής** της, δηλαδή του συνόλου των όρων που αποτελούν πραγματώσεις των βιοϊατρικών εννοιών και απαρτίζουν το λεξιλόγιό της (ορολόγιο).

Το **ορολόγιο** του ΙΑΤΡΟΛΕΞΗ καταρτίστηκε κατόπιν επεξεργασίας των συλλεχθέντων βιοϊατρικών κειμένων και άλλων λεξικογραφικών πηγών ώστε να εξαχθούν οι όροι.

Θα παρουσιαστούν αναλυτικά:

- η συλλογή των όρων: πηγές και μέθοδοι
- η επεξεργασία των όρων: κανονικοποίηση και κωδικοποίηση
- η διαδικασία αναγνώρισης των πολυλεκτικών όρων

Πηγές: ειδικά λεξιλόγια και κείμενα

- **λεξιλόγια Βιοϊατρικών όρων**

MeSH-Hellas (έκδοση 1997)

~17.500 λήμματα (μονολεκτικοί και πολυλεκτικοί όροι)
δίγλωσσο (αγγλικά - ελληνικά)

MeSH® (έκδοση 2006)

- **Σώμα Βιοϊατρικών Κειμένων (ΣΒΚ) του ΙΑΤΡΟΛΕΞΗ: 11,5 εκ. λέξεων**
 - εντοπισμός του βασικού (εν χρήσει) λεξιλογίου της βιοϊατρικής
 - εξαγωγή των άγνωστων λέξεων

Στάδια επεξεργασίας



- **Στάδιο 1°:** επεξεργασία των κειμένων → εξαγωγή άγνωστων λέξεων
- **Στάδιο 2°:** επεξεργασία των άγνωστων λέξεων → εκκαθάριση του υλικού, εξαγωγή των λέξεων που παρουσιάζουν μεγάλη πιθανότητα να είναι όροι, ταξινόμηση και οργάνωση καταλόγων
- **Στάδιο 3°:** επεξεργασία των καταλόγων που προέκυψαν → αξιολόγηση των δεδομένων με βάση στατιστικά και γραμματικά κριτήρια
- **Στάδιο 4°:** γλωσσική επεξεργασία των εξαχθέντων βιοϊατρικών όρων → κανονικοποίηση και κωδικοποίηση
- **Στάδιο 5°:** περαιτέρω επεξεργασία των πολυλεκτικών όρων → κανόνες αναγνώρισης και κλίσης

Εξαγωγή των άγνωστων λέξεων

- Ως άγνωστες λέξεις θεωρήθηκαν όσοι λεκτικοί τύποι δεν περιέχονταν στο **Μορφολογικό Λεξικό** της NEUROSOFT
 - 90.000 λήμματα (= 1.500.000 κλιτικοί τύποι) της γενικής γλώσσας
 - λέξεις ιατρικής, ανατομίας κτλ. που ανήκουν ΚΑΙ στο γενικό λεξιλόγιο (π.χ. *φυματίωση, γρίπη, λύσσα, καρδιά, εγκέφαλος* κτλ.)
- Οι άγνωστες λέξεις που εντοπίστηκαν ανήκουν σε τρεις κατηγορίες:
 - υποψήφιοι όροι (υψηλή συχνότητα εμφάνισης σε ειδικά κείμενα)
 - κύρια ονόματα (ονόματα ανθρώπων, οργανισμών, χωρών κτλ.)
 - εσφαλμένοι τύποι (κομμένες λέξεις, διάφορα λάθη κτλ.)

Εξαγωγή υποψήφιων όρων (1)

- Σε πρώτη φάση, απομακρύνθηκαν με **αυτόματο τρόπο** από τον κατάλογο των άγνωστων λέξεων όσες είχαν αρχικό γράμμα κεφαλαίο (τοπωνύμια, ονόματα προσώπων, ιδρυμάτων, νοσοκομείων κτλ.).
- Σε δεύτερη φάση, το υλικό εξετάστηκε προσεκτικά ώστε να απομακρυνθούν με **χειρωνακτικό τρόπο** :
 - εσφαλμένοι τύποι λόγω μετατροπής των κειμένων (λέξεις κολλημένες μεταξύ τους ή κολλημένες με αριθμούς και άλλα σύμβολα, λέξεις που τους λείπουν γράμματα κ.ο.κ.)
 - γραμματικά εσφαλμένοι τύποι (ορθογραφικά λάθη, παλαιότερη ορθογραφία κ.ο.κ.): π.χ. *ωοθηλακιορρηξία, κοκκύτης*
 - δύσχρηστοι τύποι της καθαρεύουσας (λόγιες γενικές ουσιαστικών, λόγιοι ρηματικοί τύποι κ.ο.κ.)

Εξαγωγή υποψήφιων όρων (2)

Αποτελέσματα:

- 25.000 άγνωστοι λεκτικοί τύποι
→ 87% είναι όροι ή συστατικά όρων της βιοϊατρικής
- 9.500 λεκτικοί τύποι (κοινοί στο ΣΒΚ και στο MeSH)
→ 12,5 % άγνωστοι τύποι στο Μορφολογικό Λεξικό

Κατάρτιση του ορολογίου (1)

- κατάρτιση ενός αρχικού καταλόγου υποψήφιων μονολεκτικών και πολυλεκτικών όρων από τις πηγές που αναφέρθηκαν
- περαιτέρω επεξεργασία και έλεγχος για την επιστημονική τους εγκυρότητα και τη γλωσσική τους ορθότητα από τους ειδικούς επιστήμονες του έργου, γιατρούς και γλωσσολόγους
- έλεγχος στο ΣΒΚ των όρων που εντοπίστηκαν στο MeSH μέσω του Συλλογέα λεξιλογικών συνάψεων (**CONCORDANCER**)

Κατάρτιση του ορολογίου (2)

Αποκλείστηκαν από το ορολόγιο του ΙΑΤΡΟΛΕΞΗ:

- όροι που εντοπίστηκαν στο MeSH αλλά δεν εμφανίστηκαν καθόλου στα κείμενα του ΣΒΚ και του Google.gr
- όροι λατινικής προέλευσης ή με ξένο όνομα το οποίο δεν έχει αποδοθεί στα ελληνικά, π.χ. *Coxiella*, *biomphalaria*, *bulinus*
 - Εξαιρούνται τα αρκτικόλεξα και τα ακρώνυμα, π.χ. *DNA*, *RNA*, *AIDS*.
- κύρια ονόματα
 - Εξαιρούνται τα ξένα κύρια ονόματα που αποτελούν συστατικό πολυλεκτικού όρου, π.χ. *σύνδρομο Caplan*, *ρίκνωση Dupuytren*, *όργανο του Corti*, *φαινόμενο Shwartzman*.

Κανονικοποίηση των όρων

- Μετατροπή του πληθυντικού αριθμού σε ενικό
 - τριχοστοματίδες → τριχοστοματίδα, μυξομύκητες → μυξομύκητας*
 - αδενοβλάστες → αδενοβλάστη, ανοσοβλάστες → ανοσοβλάστης*
 - αντισώματα ηπατίτιδας Β → αντίσωμα ηπατίτιδας Β*
- Αλλαγή της σειράς της κεφαλής και των συμπληρωμάτων που λειτουργούν ως προσδιοριστικά, καθώς επίσης και των παραλλαγών του όρου με τα κόμματα
 - ινωδογόνα μη φυσιολογικά → μη φυσιολογικά ινωδογόνα*
 - αδενοϊών πρώιμες πρωτεΐνες → πρώιμες πρωτεΐνες αδενοϊών*
 - πρωτεΐνες αυγού, διαιτητικές → διαιτητικές πρωτεΐνες αυγού*

Κωδικοποίηση των όρων

Προσθήκη μορφολογικής πληροφορίας (γραμματική κατηγορία, κλίση και χαρακτηρισμό «βιοϊατρικό») και ενσωμάτωση στο Μορφολογικό Λεξικό της NEUROSOFT

- Στο πέρας του 1^{ου} χρόνου έχουν κωδικοποιηθεί **7.250** λέξεις του βιοϊατρικού τομέα
 - 4.574 ουσιαστικά
 - 2.542 επίθετα
 - 55 ρήματα (π.χ. *ωοθετώ, φωσφορυλιώνω, οστεοποιούμαι, ...*)
 - 62 μετοχές (π.χ. *σπληνεκτομηθείς, ραγείς, υποτροπιάζων, ...*)
 - 7 επιρρήματα (π.χ. *μικροσκοπικά, νευροαπεικονιστικά, παρακαρδιακά, ...*)

Συλλογή των πολυλεκτικών όρων

Για τη συγκέντρωση των πολυλεκτικών (σύμπλοκων) όρων της βιοϊατρικής ακολουθήθηκαν οι εξής διαδικασίες:

- **Επεξεργασία του MeSH**
 - για να εντοπιστεί το βασικό λεξιλόγιο της βιοϊατρικής ορολογίας που εκφράζεται μέσω πολυλεκτικών δομών
 - για να συγκεντρωθεί ένα ικανό και επαρκές σώμα όρων το οποίο θα χρησιμεύσει ως βάση για τον προσδιορισμό μορφοσυντακτικών σχημάτων (morphosyntactic patterns)
- **Έλεγχος των πολυλεκτικών όρων του MeSH στο ΣΒΚ**
 - για να επιβεβαιωθούν τα μορφοσυντακτικά σχήματα που είχαν καταγραφεί
 - για να αναδειχθεί η μορφολογική και συντακτική τους ποικιλία
- **Μελέτη και επεξεργασία των ~ 5.000 εξαχθέντων πολυλεκτικών όρων**

Ορισμός του πολυλεκτικού σύμπλοκου όρου

Στις ειδικές γλώσσες, **πολυλεκτικός σύμπλοκος όρος** ονομάζεται κάθε όρος που αποτελείται από δύο ή περισσότερες λέξεις, τα **οροστοιχεία** του, και ο οποίος αντιπροσωπεύει μία αντίστοιχη έννοια, ειδική ή γενική, του σχετικού θεματικού πεδίου.

Ο πολυλεκτικός όρος αποτελεί ένα δυαδικό σχηματισμό που μπορεί να αναλυθεί σε δύο μέρη, τα άμεσα συνθετικά, τα οποία αποτελούνται από ένα ή περισσότερα οροστοιχεία το καθένα. Συνήθως, το ένα από τα άμεσα συνθετικά είναι το **προσδιοριζόμενο** και το άλλο το **προσδιοριστικό**.

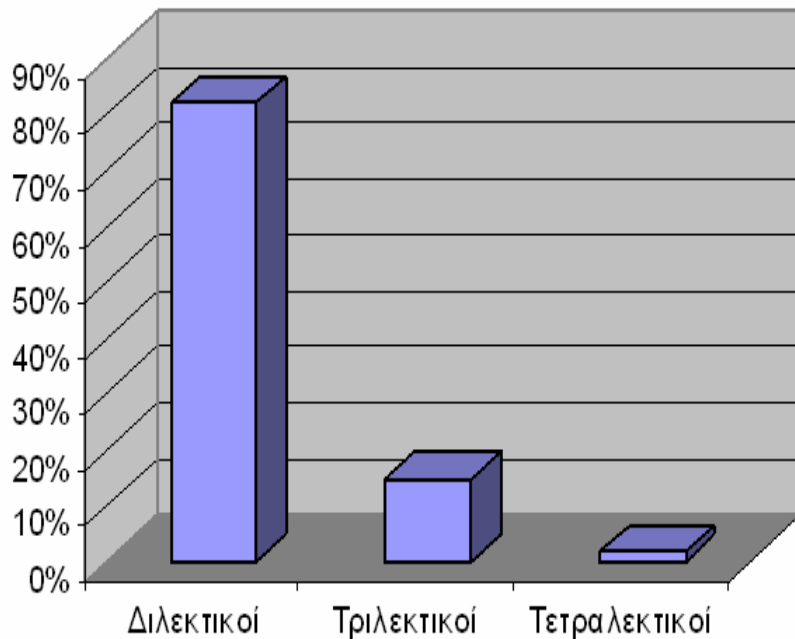
Οι πολυλεκτικοί όροι που μελετήσαμε είναι ονοματικές φράσεις, π.χ. *νωτιαίος μυελός*, *ιός της λύσσας*, που αντιστοιχούν σε καθιερωμένες έννοιες του βιοϊατρικού τομέα.

Αναγνώριση πολυλεκτικών όρων

Το σύστημα αναγνώρισης και κλίσης των πολυλεκτικών όρων του ΙΑΤΡΟΛΕΞΗ βασίστηκε στην ανάλυση ενός σώματος ~ 5.000 βιοϊατρικών πολυλεκτικών όρων που αντλήθηκε από διαφορετικές σημασιολογικές κατηγορίες της αρχικής ταξινόμιας.

- περιγραφή της συντακτικής τους δομής (~200 διαφορετικές δομές)
- διατύπωση μορφοσυντακτικών κανόνων (19 κανόνες)
- καταγραφή των οροστοιχείων τους (λεξικά και μη λεξικά στοιχεία)
- αποτύπωση της ορολογικής τους σύστασης (συνδυασμός όρων και μη όρων)
- προσθήκη σημασιολογικής πληροφορίας (σχήματα συνεμφάνισης σημασιολογικών κατηγοριών)

Κατανομή των δομών



- διάκριση σε δομές, ανάλογα με τον αριθμό των οροστοιχείων
- μέτρηση του αριθμού των αντίστοιχων όρων
 - Το 82% των πολυλεκτικών όρων είναι διλεκτικοί
 - Το 15% τριλεκτικοί
 - Το 1,9% τετραλεκτικοί
 - Το 1,1% πενταλεκτικοί και άνω
- μελέτη των πρωτογενών δομών (διλεκτικών όρων)
- μελέτη των επαυξημένων δομών (τριλεκτικών, τετραλεκτικών κτλ.)

Σχηματισμός διλεκτικών όρων

Οι διλεκτικοί όροι της βιοϊατρικής κατατάσσονται στις τέσσερις γενικές κατηγορίες πολυλεκτικών ονοματικών σύνθετων της νέας ελληνικής:

1. επίθετο+ουσιαστικό
π.χ. ρυθμιστικό γονίδιο, οφθαλμικό διάλυμα, βραχιόνιο πλέγμα
2. ουσιαστικό+ουσιαστικό σε γενική
π.χ. συνθάση χιτίνης, κοιλίες του εγκεφάλου, ιός της λύσσας
3. ουσιαστικό+πρόθεση+ουσιαστικό
π.χ. αντισυλληπτικά από το στόμα, δηλητηρίαση από σταφυλόκοκκο
4. ουσιαστικό+ουσιαστικό (ομοιόπτωτο)
π.χ. Βάκιλος μεγαθήριο, μέλος φάντασμα

Σχηματισμός τριλεκτικών και τετραλεκτικών όρων

Διεργασίες παραγωγής των επαυξημένων δομών:

- προσθήκη νέου προσδιοριστικού
νευρικό σύστημα → κεντρικό νευρικό σύστημα
λιπαρά οξέα → ακόρεστα λιπαρά οξέα
- προσθήκη νέου προσδιοριζόμενου
ερυθρά αιμοσφαίρια → όγκος ερυθρών αιμοσφαιρίων
- σύνδεση προσδιοριστικών (παρατακτική, συμπλεκτική, διαζευκτική)
άνω κοιλία, κάτω κοιλία → άνω-κάτω κοιλία
κισσοί οισοφάγου, κισσοί στομάχου → κισσοί οισοφάγου και στομάχου
μεσοκολπικό διάφραγμα, μεσοκοιλιακό διάφραγμα → μεσοκολπικό ή μεσοκοιλιακό διάφραγμα
- σύνδεση προσδιοριζόμενων
χολικά οξέα, χολικά άλατα → χολικά οξέα και άλατα
σημεία λοίμωξης, συμπτώματα λοίμωξης → σημεία και συμπτώματα λοίμωξης

Ορολογική σύσταση

Επαναπροσδιορισμός των συντακτικών δομών λαμβανομένης υπόψη της εσωτερικής τους σύστασης ως προς τη συνεμφάνιση όρων και μη όρων

Μήκος	Σχήμα	Παράδειγμα
2	<u>Ουσ+Ουσ</u>	παράγοντας συμπληρώματος
2	<u>Ουσ+ΜΟρ</u>	απώλεια δοντιών, όγκος παλμού
2	<u>ΜΟρ +Ουσ</u>	καθετήρες διαρκείας
2	<u>ΜΟρ+ ΜΟρ</u>	ιός της λύσσας, έλκος κνημών
2	<u>E+Ουσ</u>	υαλώδες σώμα, αυξητικός παράγοντας, ερυθρός πυρήνας
3	<u>E+E+Ουσ</u>	βραδεία αρνητική μεταβολή
3	<u>E+E+ΜΟρ</u>	προσωρινή μερική οδοντοστοιχία
3	<u>E+ΔΟρ</u>	κεντρικό νευρικό σύστημα, ανοικτός αρτηριακός πόρος
3	<u>ΜΟρ+ΔΟρ</u>	ιός δάγκειου πυρετού
4	<u>Ουσ+ ΤΡΟρ</u>	χρόνος πήξης ολικού αίματος, παρόξυνση <u>Χ.Α.Π.</u>
4	<u>E+ ΤΡΟρ</u>	<u>ανοσαβλαστικό</u> λέμφωμα μεγάλων κυττάρων
4	<u>ΤΡΟρ+ ΜΟρ</u>	λείες μυϊκές ίνες αγγείων
4	<u>ΜΟρ+ ΤΡΟρ</u>	σύνδρομο απόσυρσης τοξικής ουσίας
4	<u>ΔΟρ+ ΔΟρ</u>	υποφυσιακές ορμόνες οπίσθιου λοβού, φακοί επαφής συνεχούς χρήσης

Κατηγορίες οροστοιχείων

1. Λεξικά στοιχεία

- λόγιοι τύποι (π.χ. *έσω ους, ιερό οστούν*)
- ελληνικά ακρώνυμα (π.χ. *Χ.Α.Π., Ο.Μ.Λ.*)
- λέξεις με λατινικούς χαρακτήρες (π.χ. *αγγειοπλαστική με laser*)
- ονόματα με λατινικούς χαρακτήρες (π.χ. *σύνδρομο Down*)
- ακρώνυμα με λατινικούς χαρακτήρες (π.χ. *πρόδρομοι RNA*)

2. Μη λεξικά στοιχεία

- αριθμοί (π.χ. *ιντερλευκίνη 1, ημιχολίνιο 3*)
- ελληνικά και λατινικά γράμματα (π.χ. *Βιταμίνη C, ακτίνες Βήτα*)
- σημεία στίξης, όπως παύλες, κόμματα, παρενθέσεις (π.χ. *1,4-άλφα-γλυκοσιδάση γλυκάνης*)
- συνδυασμός αριθμών και γραμμάτων (π.χ. *ωμέγα-3 λιπαρά οξέα, κύτταρα 3T3*)
- συνδυασμός μη λεξικών και λεξικών στοιχείων (π.χ. *N-αποθεμελιωτικές οξειδοορεδουκτάσες, 2-αμινοαδιπικό οξύ, κυκλικά P-οξειδία*).

Μορφολογική ποικιλία (1)

ποικιλία σε επίπεδο γραφής:

χρήση/απουσία παύλας, πεζά/κεφαλαία, αναπαράσταση ελληνικών γραμμάτων, αναπαράσταση αριθμητικών, μεταγραφή ξένων ονομάτων

π.χ.

ωμέγα-3 λιπαρά οξέα / *ωμέγα 3* λιπαρά οξέα

1,4-άλφα-γλυκοσιδάση γλυκάνης / *1, 4 άλφα γλυκοζιδάση* γλυκάνης

β-αλανίνη / *βήτα αλανίνη*

α 1-αντιθρυψίνη / *άλφα 1-αντιθρυψίνη*

enterococcus faecium / *εντερόκοκκος faecium*

διοξυγονάση / *διοξυγενάση 4-υδροξυφαινυλοπυροσταφυλικού*

κτλ.

Μορφολογική ποικιλία (2)

ποικιλία σε επίπεδο ορθογραφίας:

παλαιότερη/νεότερη ορθογραφία, παράλληλοι τύποι

Π.Χ.

μπορέλια burgdorferi / μπορρέλια burgdorferi

βάκιλος άνθρακα / Βάκιλλος άνθρακα

διοξείδιο του θείου / διοξίδιο του θείου

κτλ.

Μορφολογική ποικιλία (3)

ποικιλία σε μορφοσυντακτικό επίπεδο:

παραγωγή, σύνθεση, σχηματισμός ελληνικών και ξένων ακρώνυμων, λόγιοι τύποι, αλλαγή πτώσης, αλλαγή γραμματικής κατηγορίας

π.χ.

χρόνια αναπνευστική πνευμονοπάθεια / Χ.Α.Π.

υποδοχείς ανοσοσφαιρίνης E / υποδοχείς IgE

λιποπρωτεΐνες υψηλής πυκνότητας / λιποπρωτεΐνες HDL

διοξείδιο του άνθρακα / διοξείδιο του άνθρακος

αρτηρίες του εγκεφάλου / εγκεφαλικές αρτηρίες

πρωτεϊνοκινάση C / πρωτεϊνική κινάση C

κτλ.

Σημασιολογική ανάλυση

Στάδιο 1^ο: σημασιολογική κατηγοριοποίηση των όρων

Η γνώση των σημασιολογικών κατηγοριών στις οποίες ανήκουν οι έννοιες, και κατ' επέκταση οι όροι που τις κατασημαίνουν, αντλήθηκε από την **οντολογία** του ΙΑΤΡΟΛΕΞΗ, η οποία στηρίχτηκε στο Μεταθησαυρό του UMLS.

Για τα συνθετικά των όρων, κυρίως επίθετα, που δεν περιλαμβάνονται ως αυτόνομες έννοιες ούτε στην οντολογία του ΙΑΤΡΟΛΕΞΗ ούτε στο Μεταθησαυρό του UMLS χρειάστηκε να δημιουργηθούν **νέες σημασιολογικές κατηγορίες**.

Στάδιο 2^ο: προσδιορισμός των σχημάτων συνεμφάνισης των σημασιολογικών κατηγοριών

Συγκεκριμένες υποκατηγορίες εννοιών συνδυάζονται σε συγκεκριμένα **σχήματα συνεμφάνισης** ώστε να σχηματιστούν οι πολυλεκτικοί όροι και αντίστοιχα η σημασιολογική τους κατηγορία.

Σημσιολογική κατηγορία προσδιοριστικού	Παράδειγμα
Ασθένεια, διαταραχή ή σύνδρομο	<u>νεφρωσικό</u> σύνδρομο , οφθαλμική υπόταση
Ανατομική ανωμαλία	ανεύρυσμα αορτής
Σημσιολογική κατηγορία προσδιοριστικού	
<i>Μέρος του σώματος (Μσώματος)</i>	δερματικό νόσημα, αδένας μαστού
<i>Τόπος ή θέση</i>	οπίσθιο γάγγλιο
<i>ΘέσηΜσώματος</i>	μεσοκοιλιακό διάφραγμα
<i>Λειτουργία</i>	αιμοφόρα αγγεία
<i>Σχήμα</i>	αστεροειδές γάγγλιο
<i>Υφή</i>	ινώδης χιτώνας
<i>Μέγεθος</i>	μεγάλος μαστός
<i>Ηλικία</i>	νεανικός μαστός
<i>Φύλο</i>	ανδρικά γεννητικά όργανα
<i>Ιδιότητα</i>	εμπύρετη λοίμωξη
<i>Βαθμός</i>	οξεία λοίμωξη
<i>Σειρά</i>	τρίτος γομφίος
<i>Αιτία</i>	ρευματικός πυρετός, λοίμωξη από σαλμονέλα
<i>Άτομο ή ομάδα</i>	νόσος των πτηνών , νοσήματα αγροτών
<i>Οργανική λειτουργία</i>	διατροφή βρέφους
<i>Ιστός</i>	διαταραχές κινητικότητας κροσσών επιθηλίου
<i>Ιός</i>	λοιμώξεις από ρεοϊούς
<i>Μύκητας</i>	λοιμώξεις από μικρόσπορα
<i>Βακτήριο</i>	λοιμώξεις από μπορέλια

Πίνακας 1: Σημσιολογικές κατηγορίες προσδιοριζόμενων και προσδιοριστικών συνθετικών

Μορφολογική δομή	Σημασιολογική δομή	Παράδειγμα
E+Ουσ	Μερίδιο ή Ιστός ή Κύτταρο ή Γονίδιο+Ασθένεια	μυοσκελετικά νοσήματα, οισοφαγική <u>κατακλίση</u> , κοιλιακό ανευρύσμα κυτταρική νέκρωση
	ΘέσηΜερίδιο+Ασθένεια	επιχειλίου έρπητας
	Αιτία+Ασθένεια	αλλεργική επιπεφυκίτιδα, διαβητικό κώμα, ρευματική καρδιοπάθεια, νεφρογενής υπέρταση
	Ιδιότητα+Ασθένεια	εκφυλιστική αρθρίτιδα ανιάτη νόσος
Ουσ+ΟυσΓεν	Ασθένεια+Μερίδιο	νοσήματα εγκεφάλου
	Αναπλασματική+Μερίδιο	κύστη παγκρέατος πρόπτωση ορθού στένωση τραχείας
	Ασθένεια+Ιστός ή Κύτταρο ή Γονίδιο	νοσήματα αίματος
	Ασθένεια+Οργανική λειτουργία	νοσήματα μεταβολισμού επιπλακές εγκυμοσύνης
	Ασθένεια+Άτομο ή Ομάδα	νοσήματα αγροτών νοσήματα αιγοειδών
Ουσ+Όνομα	Ασθένεια ή Σύνδρομο+Όνομα	νόσος <u>Wolman</u> , νόσος <u>Crohn</u> , σύνδρομο <u>Chiari-Frommel</u> , σύνδρομο <u>Dow</u>
E+E+Ουσ	Αιτία+Ασθένεια	ιογενή δερματικά νοσήματα, μυκητιασικά πνευμονικά νοσήματα
	Σύμπτωμα ή Βλάβη +Ασθένεια	εκζεματικά δερματικά νοσήματα, αποφρακτικά πνευμονικά νοσήματα
	Ιδιότητα+Ασθένεια	αυτοάνοσα δερματικά νοσήματα
Ουσ+ΠΡΟΘαπό+ΟυσΑπ	Ασθένεια+Ιός ή Μύκητας ή Βακτήριο ή Χλαμύδια	λοιμώξεις από <u>αιρεόπλάσμα</u> , μηνιγγίτιδα από <u>λίστέρια</u> , λοιμώξεις από σαλμονέλα
Ουσ+ΕΠΙΡύπου+Σύμβολο	Ασθένεια+Σύμβολο	διαβήτης τύπου I, <u>υπερλιπιδαιμία</u> τύπου V

Πίνακας 2: Σημασιολογικές δομές συνεμφάνισης προσδιοριζόμενων και προσδιοριστικών συνθετικών

για τους όρους της κατηγορίας «Ασθένεια, διαταραχή ή σύνδρομο»

Εργασίες σε εξέλιξη



- **Μορφολογικό Λεξικό: κωδικοποίηση και προσθήκες**
 - Στόχος: ~ 10.000 βιοϊατρικούς όρους
 - Μονολεκτικοί όροι και συνθετικά πολυλεκτικών όρων
- **Αναγνώριση των πολυλεκτικών όρων**
 - Έλεγχος των κανόνων αναγνώρισης και κλίσης με βάση τα εξαγόμενα αποτελέσματα
- **Οντολογία του ΙΑΤΡΟΛΕΞΗ**
 - Προσθήκη σημασιολογικής κατηγορίας στους όρους που έχουν εξαχθεί από το ΣΒΚ και ένταξή τους στην οντολογία