

# ΙΑΤΡΟΛΕΞΗ



## Όροι της βιοϊατρικής: συλλογή και επεξεργασία του γλωσσικού υλικού

Μαρίνα Πανταζάρα, Δρ. Γλωσσολογίας

# 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 βιοϊατρικών όρων: πηγές και μέθοδοι
- η επεξεργασία των βιοϊατρικών όρων: κανονικοποίηση και κωδικοποίηση
- η μεθοδολογία αναγνώρισης των πολυλεκτικών όρων

# Συλλογή των ελληνικών όρων της Βιοϊατρικής



**ΠΗΓΕΣ:** ειδικά λεξιλόγια και κείμενα

- **λεξιλόγια Βιοϊατρικών όρων**
  - MeSH-Hellas** (ελληνική έκδοση 1997, έντυπη μορφή)
    - ~17.500 λήμματα (μονολεκτικοί και πολυλεκτικοί όροι)
    - δίγλωσσο (αγγλικά - ελληνικά)
  - MeSH®** (αγγλική έκδοση 2006)
- **Σώμα Βιοϊατρικών Κειμένων (ΣΒΚ) του ΙΑΤΡΟΛΕΞΗ: 11,5 εκ. λέξεων**
  - εντοπισμός του βασικού εν χρήσει λεξιλογίου της Βιοϊατρικής

# Στάδια επεξεργασίας



- **Στάδιο 1<sup>ο</sup>:** επεξεργασία των **κειμένων**
  - εξαγωγή άγνωστων λέξεων και άρα υποψήφιων όρων
- **Στάδιο 2<sup>ο</sup>:** επεξεργασία των **άγνωστων λέξεων**
  - εκκαθάριση του υλικού, εξαγωγή των λέξεων που παρουσιάζουν μεγάλη πιθανότητα να είναι όροι, ταξινόμηση και οργάνωση καταλόγων
- **Στάδιο 3<sup>ο</sup>:** επεξεργασία των **καταλόγων** που προέκυψαν
  - αξιολόγηση των δεδομένων με βάση στατιστικά και γραμματικά κριτήρια
- **Στάδιο 4<sup>ο</sup>:** γλωσσική επεξεργασία των **εξαχθέντων βιοϊατρικών όρων**
  - κανονικοποίηση και κωδικοποίηση
- **Στάδιο 5<sup>ο</sup>:** εξειδικευμένη 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**:
  - ✓ ασφαμένοι τύποι λόγω μετατροπής των κειμένων (π.χ. λέξεις κολλημένες μεταξύ τους ή κολλημένες με αριθμούς και άλλα σύμβολα, λέξεις που τους λείπουν γράμματα κ.ο.κ.)
  - ✓ γραμματικά ασφαμένοι τύποι (ορθογραφικά λάθη, παλαιότερη ορθογραφία κ.ο.κ.):  
π.χ. *ωθηλακιορρηξία, κοκκύτης, λόξυγγας, λύγγας, γεντιοβίωση...*
  - ✓ δύσχρηστοι τύποι της καθαρεύουσας (π.χ. λόγιοι ρηματικοί τύποι).

# Εξαγωγή υποψήφιων όρων (2)



## Αποτελέσματα:

25.000 άγνωστοι λεκτικοί τύποι

→ 21.750 είναι όροι ή συστατικά όρων της βιοϊατρικής

# 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
- μονολεκτικοί 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 Carlan*, *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 χαρακτηρισμός «βιοϊατρικό»
- Ενσωμάτωση στο Μορφολογικό Λεξικό της NEUROSOFT
- Μέχρι στιγμής, έχουν κωδικοποιηθεί περισσότερες από **9.000** λέξεις του Βιοϊατρικού τομέα
  - 6.400 ουσιαστικά
  - 2.600 επίθετα
  - 55 ρήματα (π.χ. *ωοθετώ, φωσφορυλιώνω, οστεοποιούμαι, ...*)
  - 60 μετοχές (π.χ. *σπληνεκτομηθείς, ραγείς, υποτροπιάζων, ...*)
  - 10 επιρρήματα (π.χ. *μικροσκοπικά, νευροαπεικονιστικά, παρακαρδιακά, ...*)

# Συλλογή των πολυλεκτικών όρων της βιοϊατρικής



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 ~ 6.000 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**, κτλ.) που αντιστοιχούν σε καθιερωμένες **έννοιες** του βιοϊατρικού τομέα.

# Σημασιολογικές κατηγορίες Βιοϊατρικών πολυλεκτικών όρων



Το σύστημα αναγνώρισης και κλίσης των πολυλεκτικών όρων του ΙΑΤΡΟΛΕΞΗ βασίστηκε στην ανάλυση ενός σώματος ~ 6.000 βιοϊατρικών πολυλεκτικών όρων που αντλήθηκε από διαφορετικές σηματολογικές κατηγορίες της αρχικής ταξινόμησης.

ΣΗΜΑΣΙΟΛΟΓΙΚΗ ΚΑΤΗΓΟΡΙΑ	ΠΑΡΑΔΕΙΓΜΑ
ΓΟΝΙΔΙΑ:	γονίδιο <i>rol</i> , δομικά γονίδια εντόμων
ΑΣΘΕΝΕΙΕΣ:	πνευμονικό οίδημα, μηνιγγίτιδα από λιστέρια
ΣΥΝΔΡΟΜΑ:	σύνδρομο <i>Munchausen</i> , νεφρωσικό σύνδρομο
ΟΡΓΑΝΑ & ΜΕΡΗ ΣΩΜΑΤΟΣ:	ιγνυακή αρτηρία, κοιλιακό τοίχωμα
ΣΥΜΠΤΩΜΑΤΑ:	παρεγκεφαλιδική αταξία, βράγχος φωνής
ΙΟΙ:	ιός ηπατίτιδας Β, συγκυτιακοί ιοί αναπνευστικού
ΧΗΜΙΚΕΣ & ΦΑΡΜΑΚΕΥΤΙΚΕΣ ΟΥΣΙΕΣ:	ισότοπο ιωδίου, εμβόλιο AIDS
ΔΙΑΔΙΚΑΣΙΕΣ ΘΕΡΑΠΕΙΑΣ & ΠΡΟΛΗΨΗΣ:	πλύση στομάχου, προληπτική ορθοδοντική

# Μεθοδολογία ανάλυσης των πολυλεκτικών όρων



## Μορφοσυντακτικό επίπεδο

1. περιγραφή της συντακτικής τους δομής
  - εντοπισμός των γενικών μορφοσυντακτικών δομών σχηματισμού των πολυλεκτικών όρων (απλών και επαυξημένων)
  - διατύπωση μορφοσυντακτικών κανόνων (19 κανόνες)
2. ορολογική ανάλυση → συνδυασμός όρων και μη όρων

## Λεξιλογικό επίπεδο

3. προσδιορισμός των στοιχείων που τους απαρτίζουν → συνδυασμός λεξικών και μη λεξικών στοιχείων

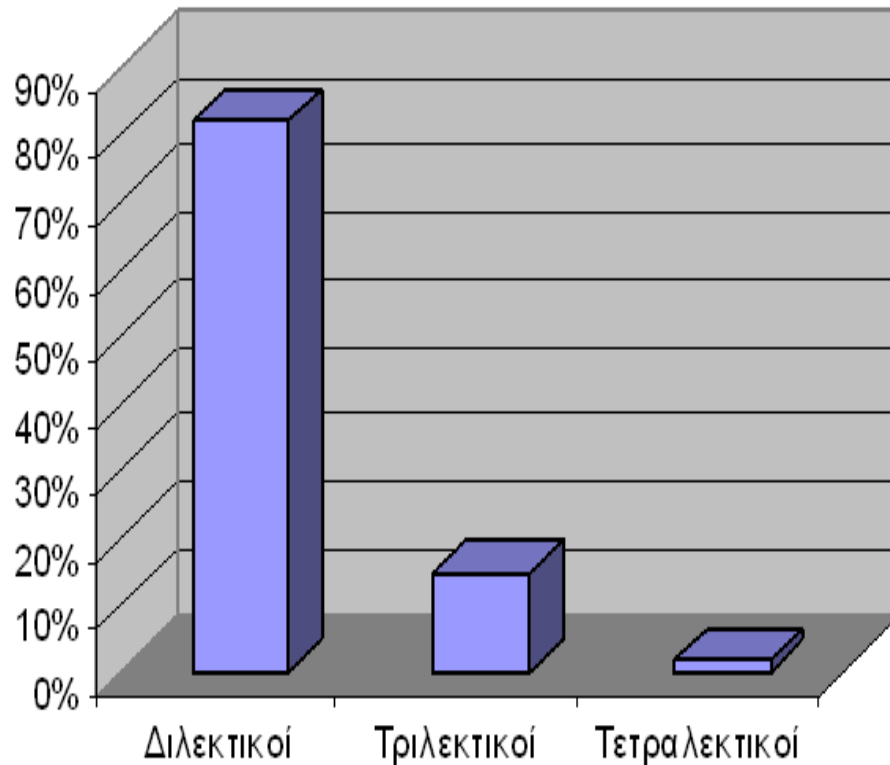
## Μορφολογικό επίπεδο

4. καταγραφή της μορφολογικής ποικιλίας → εντοπισμός των παραλλαγών ενός όρου

## Σημασιολογικό επίπεδο

5. περιγραφή των σημασιολογικών κατηγοριών στις οποίες ανήκουν τα οροστοιχεία των όρων → εντοπισμός των σχημάτων συνεμφάνισης για ορισμένες σημασιολογικές κατηγορίες πολυλεκτικών όρων

# Κατανομή των μορφοσυντακτικών δομών των πολυλεκτικών όρων



- Το **82%** των πολυλεκτικών όρων είναι *διλεκτικοί*
- Το **15%** *τριλεκτικοί*
- Το **1,9%** *τετραλεκτικοί*
- Το **1,1%** *πενταλεκτικοί* και άνω



# Σχηματισμός διλεκτικών όρων (1)



Οι διλεκτικοί όροι που μελετήθηκαν κατατάσσονται στις τέσσερις γενικές κατηγορίες πολυλεκτικών ονοματικών συνθέτων της Νέας Ελληνικής:

❖ επίθετο+ουσιαστικό:

*βραχιόνιο πλέγμα, ερυθρά αιμοσφαίρια, μεταδοτικό νόσημα*

❖ ουσιαστικό+ουσιαστικό σε γενική:

*συνθάση χιτίνης, κοιλίες του εγκεφάλου, ιός της λύσσας*

❖ ουσιαστικό+πρόθεση+ουσιαστικό:

*αντισυλληπτικά από το στόμα, δηλητηρίαση από σταφυλόκοκκο*

❖ ουσιαστικό+ουσιαστικό (ομοιόπτωτο):

*βάκιλος μεγαθήριο, μέλος φάντασμα, έντομα διαβιβαστές*

# Σχηματισμός διλεκτικών όρων (2)



## 1. Κανένα συνθετικό δεν μπορεί να προσδιοριστεί ξεχωριστά

το νευρικό σύστημα

\*το πιο νευρικό σύστημα

το ρυθμιστικό γονίδιο

\*το πολύ ρυθμιστικό γονίδιο

## 2. Μη δυνατή η σύνδεση (παρατακτική ή συμπλεκτική) ενός συνθετικού με ένα ξένο στοιχείο

άμεσες καρδιακές μαλάξεις

\*άμεσες, καρδιακές μαλάξεις

το κεντρικό νευρικό σύστημα

\*το κεντρικό και νευρικό σύστημα

## 3. Περιορισμοί ως προς την παρουσία / απουσία άρθρων

δηλητηρίαση από σταφυλόκοκκο

\*δηλητηρίαση από το σταφυλόκοκκο

## 4. Περιορισμοί ως προς τον αριθμό

νεογνικός ίκτερος

\*νεογνικοί ίκτεροι

δίαυλοι καλίου

\*δίαυλος καλίου

# Σχηματισμός διλεκτικών όρων (3)



## 5. Το προσδιοριστικό επίθετο δεν μπορεί να λειτουργήσει ως κατηγορούμενο

το ιερό οστό

\*το οστό είναι ιερό

το σύμπλεγμα βιταμινών Β

\*το σύμπλεγμα είναι βιταμινών Β

## 6. Δυνατή η απαλοιφή της κεφαλής

το νουκλεοσίδιο Q

→ το Q

το αναπνευστικό σύστημα

→ το αναπνευστικό

## 7. Μη δυνατή η αντικατάσταση ενός συνθετικού από κάποιο συνώνυμο ή αντώνυμό του

υποδοχείς αντιγόνου

\*δέκτες αντιγόνου

οδοντική εξαγωγή

\*οδοντική εισαγωγή

# Σχηματισμός διλεκτικών όρων (4)



Οι όροι σχηματίζονται με στόχο όχι μόνο την **κατασήμευση**, αλλά και την **κατηγοριοποίηση** των εννοιών:

- Δυαδική ταξινόμηση: *λευκά / ερυθρά αιμοσφαίρια*
- Πολλαπλή ταξινόμηση: *ακουστικό / οπτικό / γευστικό / οσφρητικό νεύρο*
- Μοναδική ταξινόμηση: *ομφάλιος λώρος, λοιμώδης μονοπυρήνωση*

Το προσδιοριστικό συνθετικό (εδώ, το επίθετο) προσθέτει κάποιο χαρακτηριστικό που μετατρέπει τη *γένια έννοια* του προσδιοριζόμενου συνθετικού (εδώ, του ουσιαστικού) σε *είδία έννοια* του πολυλεκτικού όρου.

# Σχηματισμός τριλεκτικών & 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 (1)



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

λύσσα	[C0034494]
ιοί	[C0521026]
ιός λύσσας	[C0034497]
νεφρική ανεπάρκεια	[C0035078]
οξεία νεφρική ανεπάρκεια	[C0022660]
χρόνια νεφρική ανεπάρκεια	[C0022661]

# Ορολογική σύσταση (2)



οδοντοστοιχίες	[C0011394]
μερική οδοντοστοιχία	[C0011460]
ολική οδοντοστοιχία	[C0011455]
άμεση ολική οδοντοστοιχία	[C0011456]
ολική οδοντοστοιχία κάτω γνάθου	[C0011457]
ολική οδοντοστοιχία άνω γνάθου	[C0011458]
γνάθος	[C0022359]
κάτω γνάθος	[C0024687]
άνω γνάθος	[C0024947]

# Ορολογική σύσταση (3)

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



# Κατηγορίες οροστοιχείων των πολυλεκτικών όρων



## 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 DNA, 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 οξειδοοδουκτάσες, 2-αμινοαδιπικό οξύ, κυκλικά P-οξειδία*)

# Μορφολογική ποικιλία των πολυλεκτικών όρων (1)



ποικιλία σε επίπεδο γραφής:

χρήση/απουσία παύλας, πεζιά/κεφαλαία, αναπαράσταση των ελληνικών γραμμάτων, αναπαράσταση των αριθμητικών, μεταγραφή ή όχι των ξένων ονομάτων

π.χ.

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

*β-αλανίνη* / *βήτα αλανίνη*

*α 1-αντιθρυψίνη* / *άλφα 1-αντιθρυψίνη*

*enterococcus faecium* / *εντερόκοκκος faecium*

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

# Μορφολογική ποικιλία των πολυλεκτικών όρων (2)



ποικιλία σε επίπεδο ορθογραφίας:

παλαιότερη/νεότερη ορθογραφία, παράλληλοι τύποι

π.χ.

*μπορέλια burgdorferi / μπορρέλια burgdorferi*

*βάκιλος άνθρακα / Βάκιλλος άνθρακα*

*διοξειδίο του θείου / διοξίδιο του θείου*

*διοξυγονάση / διοξυγενάση 4-υδροξυφαινυλοπυροσταφυλικού*

*ψευδοφυματίωση από γιερσίνια / υερσίνια*

*λεγεωνέλα / λεγιονέλα της πνευμονίας*

# Μορφολογική ποικιλία των πολυλεκτικών όρων (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*

# Σημασιολογική ανάλυση των πολυλεκτικών όρων



## Στάδιο 1<sup>ο</sup>: 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<sup>ο</sup>: 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.



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

**Πίνακας 1:** Σηματολογικές κατηγορίες προσδιοριζόμενων και προσδιοριστικών συνθετικών



Μορφοσυντακτική δομή	Σηματολογική δομή	Παράδειγμα
E+Ουσ	Μαώματος ή Ιστός ή Κύτταρο ή Γονίδιο+Ασθένεια	μυοσκελετικά νοσήματα, οισοφαγική καντιντίαση, κολλικό ανεύρυσμα κυτταρική νέκρωση
	ΘέσηΜαώματος+Ασθένεια	επιχειλίσις έρπητας
	Αιτία+Ασθένεια	αλλεργική επιπεφυκίτιδα, διαβητικό κώμα, ρευματική καρδιοπάθεια, νεφρογενής υπέρταση
	Ιδιότητα+Ασθένεια	εκφυλιστική αρθρίτιδα ανίατη νόσος
Ουσ+ΟυσΓεν	Ασθένεια+Μαώματος	νοσήματα εγκεφάλου
	Ανατ.Ανωμαλία+Μαώματος	κύστη παγκρέατος πρόπτωση ορθού στένωση τραχείας
	Ασθένεια+Ιστός ή Κύτταρο ή Γονίδιο	νοσήματα αίματος
	Ασθένεια+Οργανική λειτουργία	νοσήματα μεταβολισμού επιπλοκές εγκυμοσύνης
	Ασθένεια+Άτομο ή Ομάδα	νοσήματα αγρωτών νοσήματα αιμοειδών
Ουσ+Όνομα	Ασθένεια ή Σύνδρομο+Όνομα	νόσος Wolman, νόσος Crohn, σύνδρομο Chiari-Frommel, σύνδρομο Down
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+ΟυσΑιτ	Ασθένεια+Ιός ή Μύκητας ή Βακτήριο ή Χλαμύδια	λοιμώξεις από ουρεόπλασμα, μηνιγγίτιδα από λιστέρια, λοιμώξεις από σαλμονέλλα
Ουσ+ΕΠΙΡτύπου+Σύμβολο	Ασθένεια+Σύμβολο	διαβήτης τύπου Ι, υπερλιπιδωπρωτεϊναιμία τύπου V

**Πίνακας 2:** Σηματολογικές δομές συνεφάνισης προσδιοριζόμενων και προσδιοριστικών συνθετικών για τους άρους της κατηγορίας «Ασθένεια, διαταραχή ή σύνδρομο»

# Συμπεράσματα



Για να επιτευχθεί μεγαλύτερη ακρίβεια και αποτελεσματικότητα στην αναγνώριση των όρων του βιοϊατρικού τομέα, η γλωσσική ανάλυση θα πρέπει να εξειδικευτεί περαιτέρω, λαμβάνοντας υπόψη τα ιδιαίτερα χαρακτηριστικά των υπογλωσσών της βιοϊατρικής:

- κλινική υπογλώσσα:

*ακουστική κόπωση, σύσπαση μήτρας, ακράτεια ούρων από στρες, επώδυνος μυϊκός σπασμός, αυτοτραυματική συμπεριφορά, εξάρτηση από τοξικές ουσίες, ενδοφλέβιες ενέσεις κτλ.*

- βιομοριακή υπογλώσσα:

*ανταγωνιστές H2 υποδοχέων ισταμίνης, παράγοντας IXα, ανοσοσφαιρίνες Fc, περιοριστικά ένζυμα του DNA, ενεργοποιητές πλασμινογόνου, βήτα αμυλοειδής πρωτεΐνη, γάμα δέλτα υποδοχείς αντιγόνου T-κυττάρων κτλ.*