



ΠΑΝΕΠΙΣΤΗΜΙΟ  
ΠΕΛΟΠΟΝΝΗΣΟΥ

# ***ΥΠΟΛΟΓΙΣΤΙΚΗ ΓΛΩΣΣΟΛΟΓΙΑ***

## ***10<sup>η</sup> διάλεξη***

***Π. ΓΑΚΗΣ***

# Σύνολο

- Στη βάση των σύγχρονων Μαθηματικών (και των τυπικών γλωσσών) βρίσκεται η **θεωρία των συνόλων**
- Τα **σύνολα** (sets), λοιπόν, είναι συλλογές διακριτών αντικειμένων, είτε της πραγματικότητας είτε της φαντασίας μας ή της αφηρημένης σκέψης μας.
- Τα αντικείμενα τα οποία **περιέχει** ή από τα οποία **αποτελείται** ένα σύνολο λέγονται **στοιχεία** ή **μέλη** του συνόλου και **ανήκουν** στο σύνολο.

# Παραδείγματα συνόλου

- **Σύνολα από οντότητες πραγματικότητας:** (το σύνολο των παιχνιδιών του αδελφού μου που στο εξής θα το λέμε σύνολο  $\Pi$ . Το Star Wars Lego που αγόρασα στον αδελφό μου την Πρωτοχρονιά είναι στοιχείο ή μέλος του **συνόλου  $\Pi$** .)
- **Σύνολα με οντότητες της φαντασίας:** το σύνολο των δώδεκα θεών του Ολύμπου και θα το πούμε **σύνολο  $\Theta$** . Δηλαδή, ο Δίας ανήκει στο σύνολο  $\Theta$

# Παραδείγματα συνόλου

- **Σύνολα με οντότητες της αφηρημένης σκέψης:** είναι το σύνολο των αξιωμάτων που στηρίζουν την Ευκλείδεια Γεωμετρία και θα το πούμε **σύνολο  $E$** .

# Μαθηματική γραφή για τα σύνολα

- Όπως αυτά που λέμε στη γλώσσα τα γράφουμε κιάλας, έτσι και για τα Μαθηματικά έχουμε επινοήσει ειδική γραφή. Έχουμε λοιπόν δύο τρόπους να περιγράψουμε ένα σύνολο:

# Μαθηματική γραφή για τα σύνολα

1. Ο ένας τρόπος είναι να απαριθμήσουμε τα στοιχεία του, όπως κάνουμε για οντότητες της φαντασίας μας **για το σύνολο  $O$** .

$O = \{\Deltaίας, Αθηνά, Ποσειδών, \dots, Εστία\}$

**προσέξτε τις τρεις τελείες:** σημαίνουν μια σειρά στοιχείων του συνόλου που είναι γνωστά. Ούτε δύο, ούτε τέσσερις τελείες: τρεις, **όταν θέλουμε να δηλώσουμε συγκεκριμένη ή και άπειρη σειρά στοιχείων που είναι γνωστά.** Σημαντικό: η σειρά των στοιχείων δεν έχει καμία σημασία απολύτως.

# Μαθηματική γραφή για τα σύνολα

- Άλλη γραφή:

$$O = \{ x | x \text{ θεός του Ολύμπου} \}$$

- Χρησιμοποιούμε το γράμμα  $x$  για να συμβολίσουμε το τυχαίο στοιχείο ενός συνόλου.
- Η γραφή διαβάζεται ως εξής: «το σύνολο  $O$  απαρτίζεται από διάφορα στοιχεία  $x$  τέτοια ώστε το  $x$  είναι θεός του Ολύμπου».

- $O = \{\text{Δίας, Αθηνά, Ποσειδών, \dots, Εστία}\}$
- $O = \{x \mid x \text{ θεός του Ολύμπου}\}$

### ΔΙΕΥΚΡΙΝΙΣΗ

- αν και τα σύνολα που δίνουμε ως παραδείγματα εδώ περιέχουν στοιχεία με τουλάχιστον μία κοινή ιδιότητα, όπως π.χ. «θεός του Ολύμπου», η ύπαρξη μιας κοινής ιδιότητας δεν είναι αναγκαία για να ορίσουμε ένα σύνολο.
- Δηλαδή, εντελώς ανόμοια στοιχεία μπορούν να απαρτίζουν ένα σύνολο. Έτσι, για παράδειγμα το σύνολο

$A = \{1, \text{μουσακάς, Μπετόβεν, κλειδαριά}\}$   
είναι ένα καθόλα αποδεκτό σύνολο.



# Διαγράμματα Venn

- διαγράμματα Venn = γραφιστικός τρόπος να αναπαραστήσουμε τα σύνολα:
- Έτσι το  $\emptyset$  αναπαρίσταται ως εξής:



# Διαγράμματα Venn

- Χρησιμοποιούμε τον συμβολισμό  $x \in A$  αν το  $x$  ανήκει στο σύνολο  $A$  ή, ισοδύναμα, το  $x$  είναι μέλος ή στοιχείο του συνόλου  $A$

**Δίας  $\in O$**

- και τον συμβολισμό  $x \notin A$  αν το  $x$  δεν ανήκει στο σύνολο  $A$  ή, ισοδύναμα, δεν είναι μέλος ή στοιχείο του συνόλου  $A$

**Star Wars Lego  $\notin E$**

Σύνολο  
αξιωμάτων  
Ευκλείδεια  
Γεωμετρία

# Κενό σύνολο

- Υπάρχει ένα σύνολο που δεν περιέχει κανένα στοιχείο, το λεγόμενο κενό σύνολο (empty set).
- Το συμβολίζουμε με  $\{\}$  ή με το  $\emptyset$ .

# Υποσύνολο συνόλου

- Αν έχουμε δύο σύνολα όπου όλα τα στοιχεία του ενός, π.χ. του  $O_1$ , είναι στοιχεία του άλλου, π.χ. του  $O$ , τότε λέμε ότι το  $O_1$  είναι **υποσύνολο** (subset) του  $O$  και γράφουμε τη σχέση

$$O_1 \subseteq O$$

Για παράδειγμα, το  $O_1$  είναι το σύνολο των γυναικείων θεοτήτων του Ολύμπου, δηλαδή  $O_1 = \{\text{Ήρα, Αθηνά, Αφροδίτη, Άρτεμις, Δήμητρα, Εστία}\}$  είναι υποσύνολο του  $O$ .

## Γνήσιο υποσύνολο (proper subset)

- Εάν το  $O$  περιέχει έστω και ένα μόνο στοιχείο παραπάνω από το  $O_1$ , τότε λέμε ότι το  $O_1$  είναι **γνήσιο υποσύνολο** (proper subset) του  $O$  και γράφουμε τη σχέση:

$$O_1 \subset O$$

# Άσκηση

Ποιο από τα (1), (2) είναι σωστό και γιατί:

1.  $\emptyset \subseteq \emptyset \longrightarrow$  υποσύνολο

2.  $\emptyset \subset \emptyset \longrightarrow$  Γνήσιο υποσύνολο (περιέχει ένα ακόμη στοιχείο)

# ΑΠΑΝΤΗΣΗ

Ποιο από τα (1), (2) είναι σωστό και γιατί:

1.  $\emptyset \subseteq \emptyset$

2.  $\emptyset \subset \emptyset$

ΑΠΑΝΤΗΣΗ: 1

- Όταν 2 σύνολα είναι ίσα τότε είναι και υποσύνολο το ένα από το άλλο
- το 2 δεν ισχύει γιατί ο τελεστής "proper subset" προϋποθέτει ότι το αριστερά σύνολο έχει λιγότερα στοιχεία από το δεξιό (γνήσιο υποσύνολο) δηλ. δεν είναι ίσα τα 2 σύνολα

# Άσκηση

- $\Pi_1 = \{\text{Lego\_διαστημόπλοιο}, \text{Lego\_Darth Vader}, \text{Lego\_φρουρός}\}$
- $\Pi_2 = \{\text{Lego\_φρουρός}, \text{Lego\_διαστημόπλοιο}, \text{Lego\_Darth Vader}\}$

Ποιες προτάσεις είναι σωστές:

1.  $\Pi_1 = \Pi_2$
2.  $\Pi_1 \neq \Pi_2$
3.  $\Pi_1 \subset \Pi_2$
4.  $\Pi_1 \subseteq \Pi_2$



# ΑΠΑΝΤΗΣΗ

- $\Pi_1 = \{\text{Lego\_διαστημόπλοιο}, \text{Lego\_Darth Vader}, \text{Lego\_φρουρός}\}$
- $\Pi_2 = \{\text{Lego\_φρουρός}, \text{Lego\_διαστημόπλοιο}, \text{Lego\_Darth Vader}\}$

Ποιες προτάσεις είναι σωστές:

1.  $\Pi_1 = \Pi_2$
2.  $\Pi_1 \neq \Pi_2$
3.  $\Pi_1 \subset \Pi_2$
4.  $\Pi_1 \subseteq \Pi_2$

1 & 4

1  $\Rightarrow$  Στα σύνολα δεν έχει σημασία η σειρά των στοιχείων και αφού και τα 2 έχουν τα ίδια στοιχεία είναι ίσα (=)

4  $\Rightarrow$  Όταν 2 σύνολα είναι ίσα τότε είναι και υποσύνολο το ένα από το άλλο

# Ένωση συνόλων

- Αν έχουμε δύο σύνολα  $A$  και  $B$  μπορούμε να φτιάξουμε ένα σύνολο  $\Gamma$  τέτοιο που να περιέχει όλα τα στοιχεία του  $A$  και όλα τα στοιχεία του  $B$  και τίποτε άλλο: λέμε ότι το σύνολο  $\Gamma$  είναι η **ένωση** (union) των συνόλων  $A$  και  $B$  και γράφουμε τη σχέση:

$$A \cup B = \Gamma$$

# Άσκηση

- Είναι το  $O_3$  η ένωση των  $O_1$  και  $O_2$ ;

$O_1 = \{\text{Άρης, Ερμής}\}$

$O_2 = \{\text{Άρτεμις, Ήρα}\}$

$O_3 = \{\text{Άρης, Ερμής, Άρτεμις, Ήρα, Δίας}\}$

# Τομή συνόλων

- Αν έχουμε δύο σύνολα  $A$  και  $B$  μπορούμε να φτιάξουμε ένα σύνολο  $\Gamma$  τέτοιο που να περιέχει μόνο τα κοινά στοιχεία των  $A$  και  $B$  και τίποτε άλλο.
- λέμε ότι το σύνολο  $\Gamma$  είναι η **τομή** (intersection) των συνόλων  $A$  και  $B$  και γράφουμε τη σχέση:

$$A \cap B = \Gamma$$

# Άσκηση

- Είναι το  $O_3$  η τομή των  $O_1$  και  $O_2$ ;

$O_1 = \{\text{Άρης, Ερμής, Ήρα}\}$

$O_2 = \{\text{Άρτεμις, Ήρα, Δήμητρα}\}$

$O_3 = \{\text{Ήρα, Δίας}\}$

# Άσκηση

- Είναι το  $O_3$  η τομή των  $O_1$  και  $O_2$ ;

$O_1 = \{\text{Άρης, Ερμής, Ήρα}\}$

$O_2 = \{\text{Άρτεμις, Ήρα, Δήμητρα}\}$

$O_3 = \{\text{Ήρα, Δίας}\}$

- ΑΠΑΝΤΗΣΗ: ΟΧΙ

# Συμπερασματικά

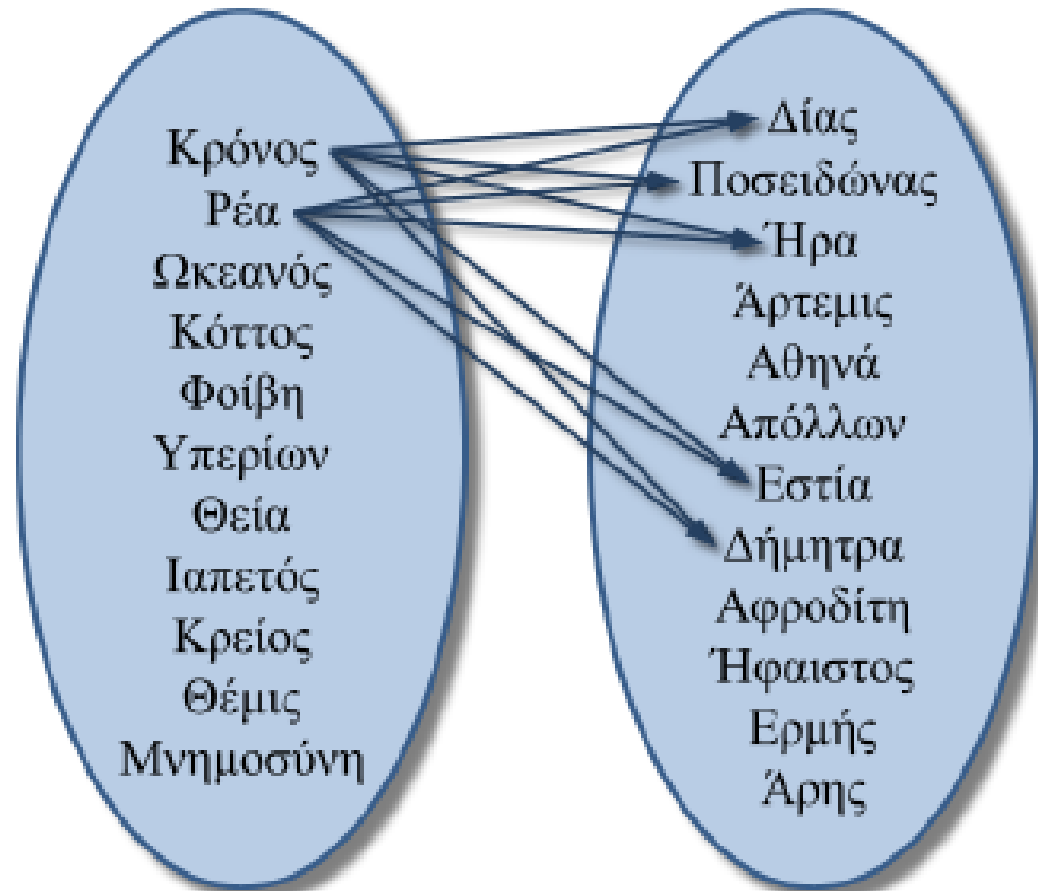
- Τα σύνολα είναι μαθηματικά αντικείμενα που θα μας χρειαστούν στη συνέχεια για να ορίσουμε τα μοντέλα αναπαράστασης της φυσικής γλώσσας.
- Άρα είναι σημαντικό να έχουμε στο μυαλό μας τι είναι αυτό το μαθηματικό αντικείμενο και τις βασικές πράξεις που γίνονται με τα σύνολα.
- Στη βάση της σκέψης μας βρίσκεται το ότι τα σύνολα είναι συλλογές αντικειμένων και τα αντικείμενα είναι μοναδικά, όπως μοναδικός είναι ο καθένας από εμάς.

# Σχέσεις

- Η έννοια της σχέσης (relation) είναι καθημερινή. Λέμε ότι ανάμεσα σε δύο σύνολα αντικειμένων υφίσταται κάποια σχέση όταν υπάρχουν καταστάσεις που για να τις περιγράψουμε χρειαζόμαστε και τα δύο.
- Για παράδειγμα, ας σκεφτούμε το σύνολο των Τιτανίδων  $T = \{\text{Κρόνος, Ρέα, Ωκεανός, Κόττος, Φοίβη, Υπερίων, Θεία, Ιαπετός, Κρείος, Θέμις, Μνημοσύνη}\}$ .
- Ο Κρόνος και η Ρέα γέννησαν τους Δία, Ήρα, Δήμητρα, Εστία, Ποσειδώνα, Πλούτωνα, που είναι στοιχεία του συνόλου  $O$ .
- Υπάρχει λοιπόν μια κατάσταση που εμπλέκει μέλη και των δύο συνόλων.
- Αυτήν τη συγκεκριμένη κατάσταση ονομάζουμε σχέση «έχει παιδί τον». Σε αυτήν την σχέση βρίσκονται ο Κρόνος και ο Δίας ως εξής: Ο Κρόνος έχει παιδί τον Δία.



# Αναπαράσταση της σχέσης «έχει παιδί τον» με διαγράμματα Venn



# Καρτεσιανό γινόμενο

- Είναι δυνατόν να απεικονίσουμε τις σχέσεις μεταξύ των στοιχείων δύο συνόλων  $A$  και  $B$  ως υποσύνολα του συνόλου που αποτελεί το **καρτεσιανό γινόμενο** (Cartesian product) των  $A$  και  $B$ .

# Καρτεσιανό γινόμενο

- Έστω, λοιπόν, ότι έχουμε δύο σύνολα, π.χ. το  $A = \{1, 2, 3\}$  και το  $B = \{\alpha, \beta\}$ .
- Το καρτεσιανό γινόμενο  $A \times B$  των συνόλων  $A$  και  $B$ , είναι το σύνολο των διατεταγμένων ζευγών όπου το πρώτο στοιχείο είναι στοιχείο του  $A$  και το δεύτερο στοιχείο του  $B$ .
- το καρτεσιανό γινόμενο των συνόλων  $A$  και  $B$  φαίνεται στη σχέση:

$$A \times B = \{ \langle 1, \alpha \rangle, \langle 1, \beta \rangle, \langle 2, \alpha \rangle, \langle 2, \beta \rangle, \langle 3, \alpha \rangle, \langle 3, \beta \rangle \}.$$

# Γραμματικές

- Η γραμματική έχει διαφορετική σημασία στην υπολογιστική γλωσσολογία.
- Δεν είναι απλώς ένα σύνολο κανόνων που περιγράφουν την **ορθή και κοινά αποδεκτή** χρήση της γλώσσας.
- Είναι ένα σύνολο κανόνων που μπορούν να δημιουργήσουν όλες τις πιθανές εκφράσεις σε μια γλώσσα.
- Μια γραμματική περιγράφει τον τρόπο σχηματισμού έγκυρης ακολουθίας συμβόλων που ανήκουν στη γλώσσα

# Γραμματικές

- Οι σύγχρονες γραμματικές έχουν χαρακτηριστικά:
  1. φραστικά συστατικών,
  2. η αναδρομή και
  3. η ανεξαρτησία συμφραζομένων

# Φραστικά Συστατικά

- μπορούμε να αναγνωρίσουμε σύνολα λέξεων τα οποία αποτελούν λειτουργικά διακριτές ομάδες μέσα στην πρόταση.
- Τις ομάδες αυτές τις αποκαλούμε **φραστικά συστατικά** (συστατικά της φράσης).
- Χρησιμοποιούμε για την αναγνώρισή τους διάφορα διαγνωστικά κριτήρια. Σ' αυτά τα συστατικά δίνονται και διακριτικά ονόματα, όπως **Ονοματική Φράση**, **Ρηματική Φράση** κ.ά.

# Παραδείγματα

Η ερωτηματική έκφραση αντικαθιστά ένα συστατικό.

*Είδα τον Γιάννη. -- Ποιόν είδες;*

Το συστατικό εδώ είναι η **Ονοματική Φράση** "τον Γιάννη".)

*Ήρθα χθές με την βροχή. -- Πότε ήρθες;*

Το συστατικό εδώ είναι η **Επιρρηματική Φράση** "χθες με την βροχή".)

# Αναδρομή

- Η αναδρομή παρουσιάζεται στη φυσική γλώσσα και, **θεωρητικά, είναι δυνατόν να υπάρξουν προτάσεις απείρου μήκους**. Αυτό δεν γίνεται στην πράξη γιατί ο άνθρωπος έχει πεπερασμένες δυνατότητες επεξεργασίας.
- Οι γραμματικές όμως που προσπαθούν να εκφράσουν τις γλωσσικές ικανότητες του ανθρώπου είναι σε θέση να παράγουν **συμβολοακολουθίες (strings) απείρου μήκους, όπως:**

*Ο Κώστας είπε ότι ο Γιάννης πιστεύει ότι ο Σπύρος νόμισε ότι η Ελένη υπεστήριξε ότι ....*



# Ανεξαρτησία Συμφραζομένων

Στην ακόλουθη μικρή *συμφραστικός ανεξάρτητη* γραμματική με *κανόνες φραστικής δομής*, τα τερματικά σύμβολα δίνονται με μικρά γράμματα και τα μη τερματικά με κεφαλαία.

S	→	NP VP
NP	→	DET (A*) N
VP	→	V (NP)
DET	→	ο, τον
N	→	σκύλος, ποντικό
V	→	νυστάζει, κυνηγά

Συχνά χρησιμοποιούμε παρενθέσεις όταν γράφουμε γραμματικούς κανόνες. Η παρένθεση εισάγει ένα είδος συντομογραφίας. Ο αστερίσκος σημαίνει καμία ή περισσότερες φορές εμφάνιση

# Ανεξαρτησία Συμφραζομένων

- ο κανόνας (1) αναλύεται στους κανόνες (1.1) και (1.2).

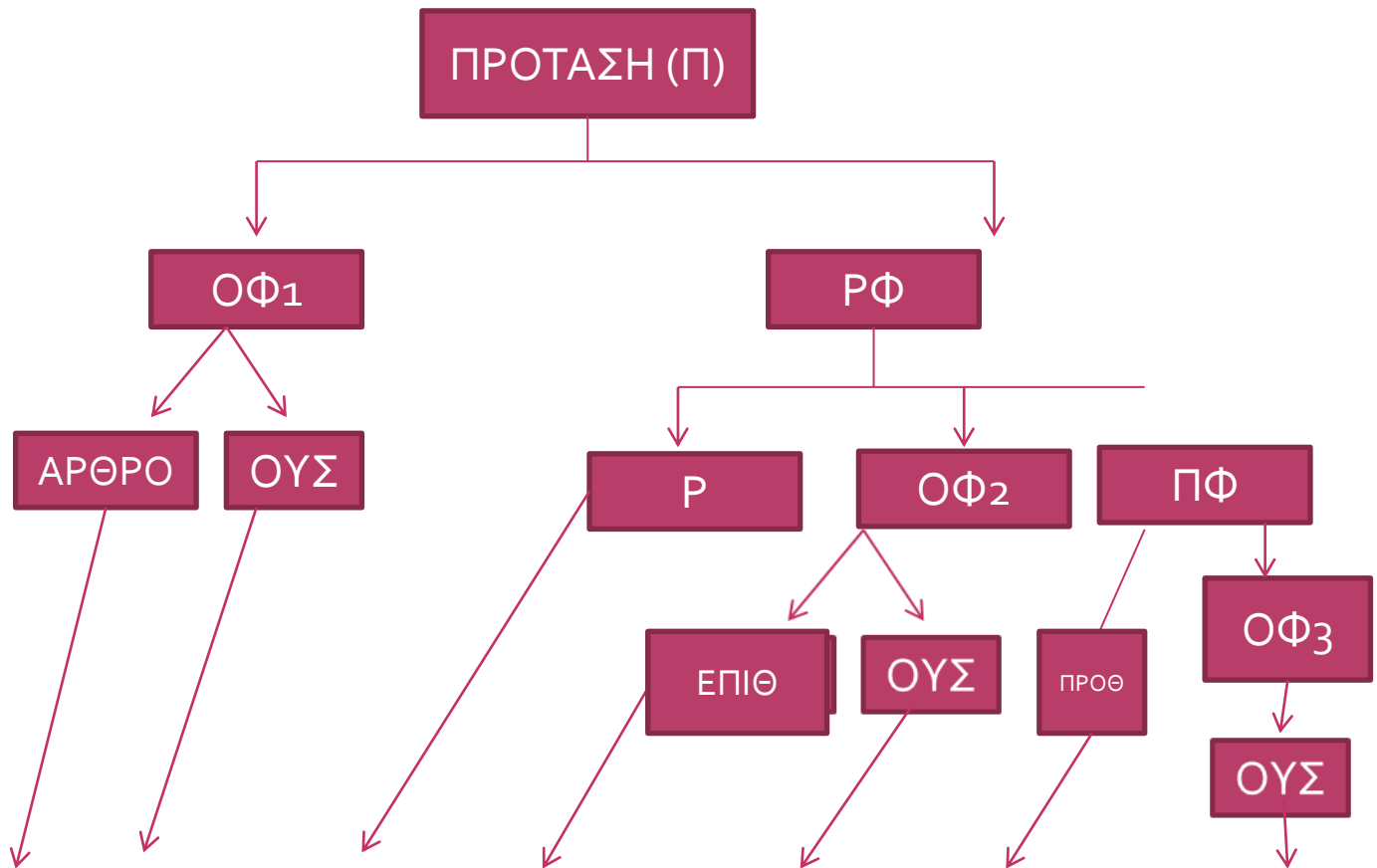
$$(1) \quad O\Phi \rightarrow AP\Theta (E) O$$

$$(1.1) \quad O\Phi \rightarrow AP\Theta E O$$

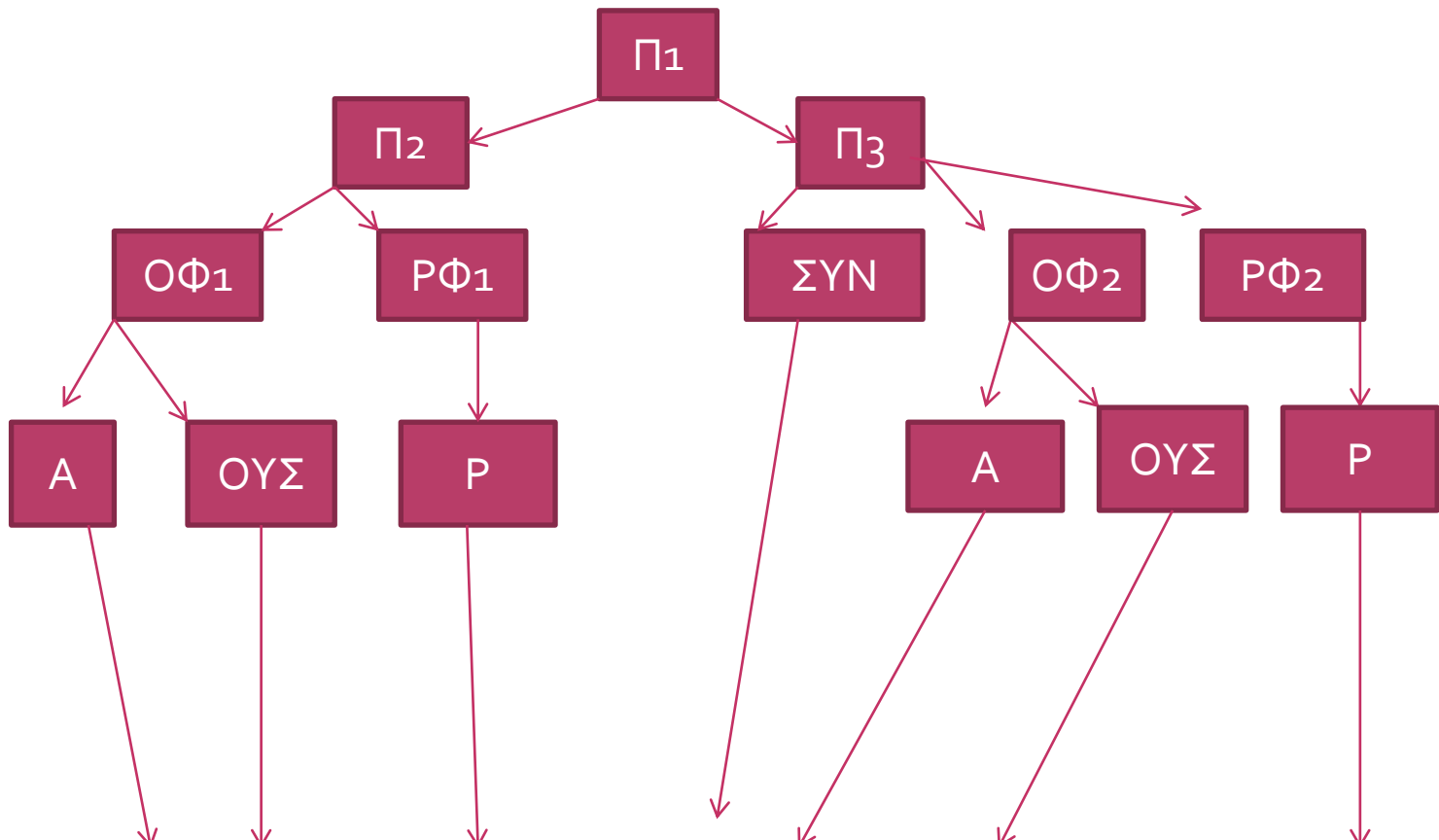
$$(1.2) \quad O\Phi \rightarrow AP\Theta O$$

# ΑΣΚΗΣΕΙΣ

- Μπορείτε να γράψετε τη γραμματική που αναγνωρίζει τις προτάσεις;
  - *ο σκύλος έδωσε ένα παξιμάδι στον ποντικό.*
  - *ο σκύλος ξέρει ότι ο ποντικός νυστάζει.*



- ο σκύλος έδωσε ένα παξιμάδι στον ποντικό.



- ο σκύλος ξέρει ότι ο ποντικός νυστάζει.

# N. Chomsky (1928 - )



# N. Chomsky

- Ο Νόαμ Τσόμσκι είναι Αμερικανός καθηγητής στο Τμήμα Γλωσσολογίας και Φιλοσοφίας του Τεχνολογικού Ινστιτούτου της Μασαχουσέτης
- Έχει συγγράψει πλήθος βιβλίων και άρθρων, ενώ έχει δώσει και εκτενείς διαλέξεις επάνω σε ένα ευρύτατο φάσμα θεμάτων τα οποία περιλαμβάνουν τη γλωσσολογία, τη φιλοσοφία και την ιστορία της διάνοησης.
- Ο Τσόμσκι είναι ο εισηγητής της λεγόμενης «γενετικής-μετασχηματιστικής γραμματικής», κυρίως με το ριζοσπαστικό γλωσσολογικό του σύγγραμμα Συντακτικές Δομές του 1957

# N. Chomsky

- Η γενετική θεωρία για τη γλώσσα, παρά τις αλλεπάλληλες εξελίξεις και αναθεωρήσεις της, χαρακτηρίζεται από κάποιες σταθερές παραδοχές όπως η ακόλουθη:

*η ικανότητα του ανθρώπου να παράγει και να κατανοεί ασύλληπτο για τον ανθρώπινο νου αριθμό προτάσεων είναι έμφυτη και μας οδηγεί βάσιμα στην υπόθεση για την ύπαρξη γλωσσικών καθολικών, γενικευμένων δομών και περιορισμών στους οποίους υπακούν όλες οι φυσικές γλώσσες, παρά την παρατηρούμενη τεράστια τυπολογική ποικιλία τους.*



# N. Chomsky

- Η ύπαρξη μιας εγγενούς «Καθολικής Γραμματικής» καθιστά το παιδί ικανό να μαθαίνει τη μητρική του γλώσσα σε ελάχιστο χρονικό διάστημα, παρά την αποσπασματικότητα των δεδομένων τα οποία προσλαμβάνει ο άνθρωπος κατά τα πρώτα χρόνια της ζωής του.
- Ο Chomsky αναφέρει ότι τα παιδιά δεν γεννιούνται «tabula rasa» (κενό χαρτί) αλλά **είναι γενετικά προδιατεθειμένα**, ώστε να μπορούν να δομήσουν την απόκτηση γνώσης. Σύμφωνα με αυτόν υπάρχει **ένας έμφυτος μηχανισμός κατάκτησης της γλώσσας** που βοηθά τα παιδιά στο να κατακτήσουν και να αποκτήσουν την γλώσσα τους.

# N. Chomsky (αποφθέγματα)

- Δεν μπορείς να εξαναγκάσεις ανθρώπους να υπακούουν πάντα δια της βίας. Γι' αυτό οι ελίτ χρησιμοποιούν «συστήματα κατήχησης» για να σιγουρέψουν, δηλαδή, με συγκατάθεση αυτό που θέλουν να κάνουν.
- Δεν μπορείς να πολεμήσεις τον κόσμο μόνος σου... Ο τρόπος για να πολεμήσεις τον κόσμο είναι μέσω της οργάνωσης...
- Δεν υπάρχει πρωτόγονη γλώσσα... Η γλώσσα είναι έμφυτη, προϊόν της φύσης μας και όχι του πολιτισμικού μας κεκτημένου.
- Ειδικός είναι κάποιος που διατυπώνει με ευφράδεια την ομοφωνία των ισχυρών.
- Η ικανότητα της ομιλίας είναι το αποτέλεσμα μιας ώριμης νοητικής κατάστασης και της αλληλεπίδρασης δύο παραγόντων: των εμφύτων ιδιοτήτων του νου και του περιβάλλοντος.
- Πηγή: <https://www.sansimera.gr/quotes/authors/355>

# Noam Chomsky

- **γλωσσική ικανότητα** (linguistic competence): αναφέρεται στη γνώση που οι ομιλητές/ακροατές έχουν για τη γλώσσα τους, το σύστημα των κανόνων που έχουν αποκτήσει ασυνείδητα έτσι ώστε να μπορούν **να παράγουν** και να **κατανοούν** άπειρο αριθμό σωστών προτάσεων και ταυτόχρονα να **αναγνωρίζουν** και να **απορρίπτουν** αυτές που δεν ανταποκρίνονται στο γραμματικό τους σύστημα.
- **γλωσσική επιτέλεση** (performance) ορίστηκε “η πραγματική χρήση της γλώσσας σε συγκεκριμένες καταστάσεις” (Chomsky 1965: 4) **το σύνολο δηλ. των γλωσσικών εκφωνημάτων που παράγονται από τον ομιλητή συμπεριλαμβανομένων πιθανών λαθών, δισταγμών εκκίνησης της ομιλίας (false starts) και άλλων προβλημάτων** που δεν έχουν σχέση με το αφηρημένο σύστημα κανόνων, αλλά μπορεί να οφείλονται σε διάφορους **μη γλωσσικούς παράγοντες** (κούραση, εκνευρισμός, κ.ά).

# Η σύσταση της γλωσσικής ικανότητας (CHOMSKY)

- Η γλωσσική ικανότητα και επομένως η γνώση της γλώσσας, κατά τον Chomsky (1980), βασίζεται σε τρεις υποκατηγορίες γνωστικών ικανοτήτων (βλ. σχήμα 3):
  - **γραμματική ικανότητα** (grammatical competence): **η ικανότητα του φυσικού ομιλητή να κατανοεί και να συνδυάζει την μορφή** (form) και τη **σημασία** (Chomsky 1980: 59)
  - **πραγματολογική ικανότητα** (pragmatic competence): η γνώση του φυσικού ομιλητή των συμβάσεων που του επιτρέπουν να **χρησιμοποιεί κατάλληλα την γλώσσα ανάλογα με την περίσταση** επικοινωνίας. Με τα λόγια του Chomsky (1980: 224 - 225) "η **πραγματολογική ικανότητα τοποθετεί τη γλώσσα στο θεσμικό περιβάλλον της χρήσης της, συσχετίζοντας προθέσεις και σκοπούς με τα υπάρχοντα γλωσσικά μέσα**".

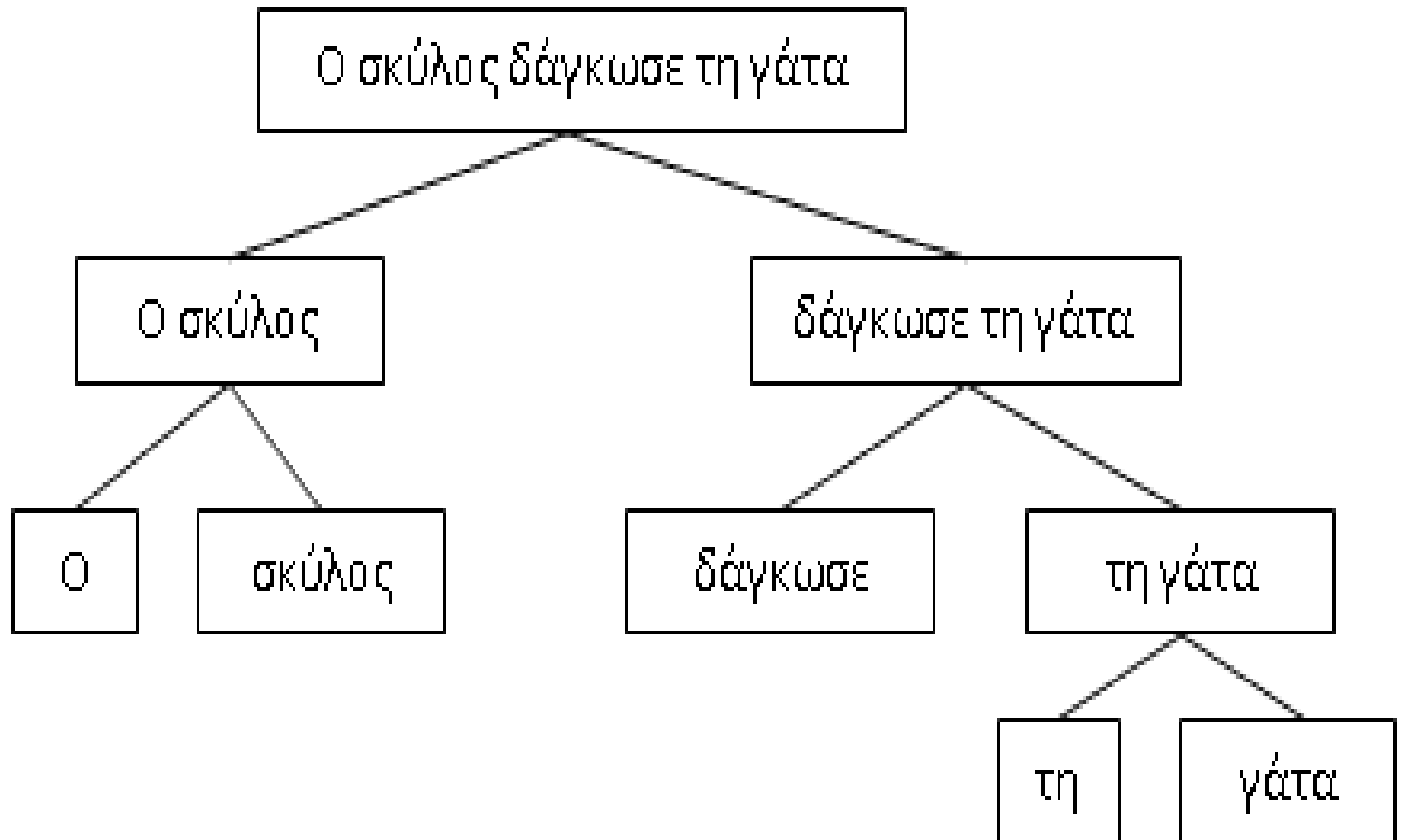
# Η σύσταση της γλωσσικής ικανότητας (CHOMSKY)

- **εννοιολογικό σύστημα** (conceptual system): η ικανότητα του φυσικού ομιλητή “να αντιλαμβάνεται, και να κατηγοριοποιεί και να συμβολίζει, ακόμα και να συλλογίζεται με ένα βασικό τρόπο” (Chomsky 1980: 55 κ.ε.). Αυτή η ικανότητα είναι περισσότερο “**πρωτόγονη**” σε σχέση με τη γραμματική ικανότητα και δεν περιορίζεται στους ανθρώπους. Ο συνδυασμός, ωστόσο, της γραμματικής και εννοιολογικής ικανότητας είναι ο καθοριστικός παράγοντας της ανάπτυξης στον άνθρωπο του πιο μοναδικού χαρακτηριστικού του, της ελεύθερης σκέψης (Chomsky 1982: 20).

# Γραμματικές N. Chomsky

- Ο N. Chomsky (Chomsky, 1956) πρότεινε αρκετούς τύπους γραμματικών διαφέρουν
  1. τόσο ως προς το είδος των κανόνων,
  2. όσο και ως προς τους τύπους εκφράσεων που παράγουν.

# Ανάλυση σε άμεσα συστατικά



# Γραμματική Φραστικής Δομής

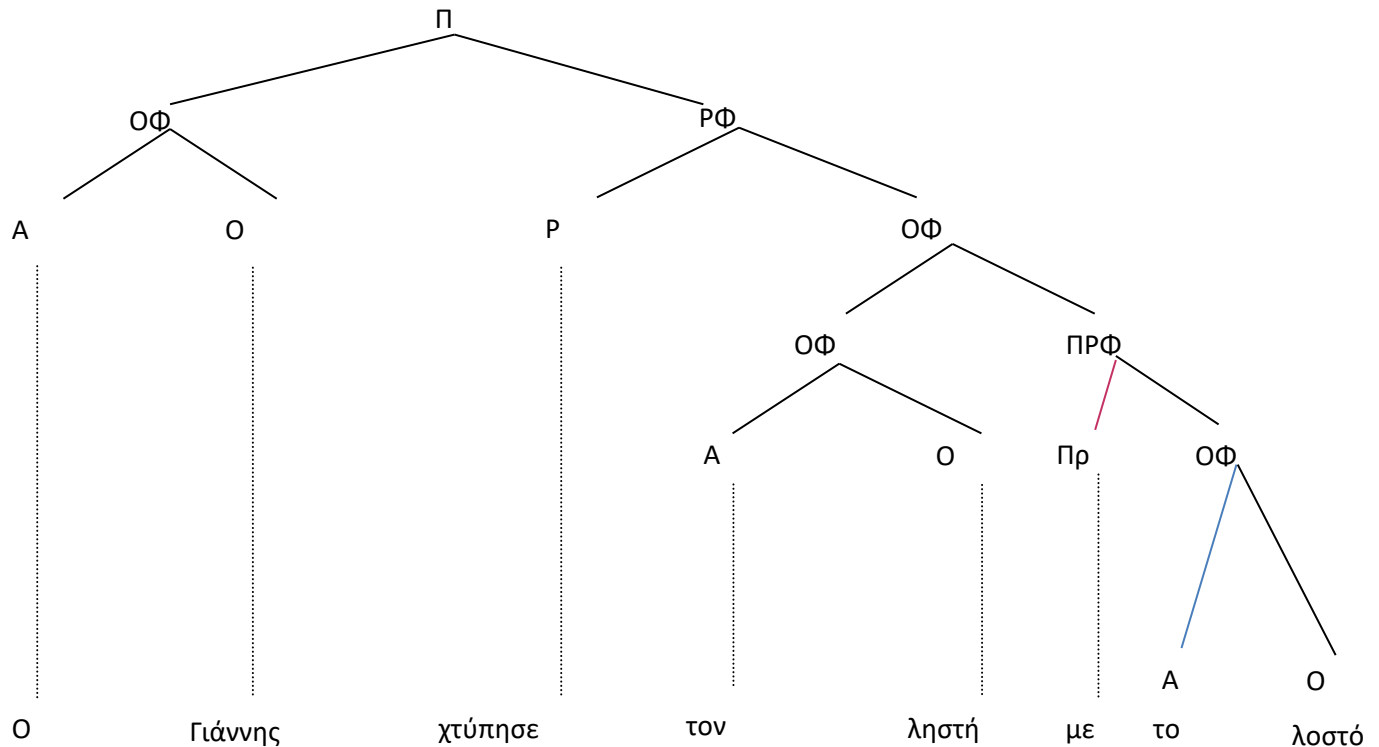
- Η ανάλυση σε άμεσα συστατικά επιτρέπει να αποτυπώσουμε **στη συντακτική μας περιγραφή** την ιεραρχική δομή που υφίσταται μεταξύ των όρων μιας πρότασης, γεγονός που δεν ήταν εφικτό με παλαιότερες παραδοσιακές μεθόδους συντακτικής ανάλυσης.
- Έτσι έγινε εφικτό να τυποποιηθούν οι συντακτικές σχέσεις και να οδηγηθούμε σε συμβολισμό που μπορεί να τις περιγράψει με γενικότητα. Το αποτέλεσμα είναι η Γραμματική Φραστικής Δομής - ΓΦΔ (Phrase Structure Grammar).
- Τα κύρια χαρακτηριστικά μιας ΓΦΔ είναι:
  1. Σύμβολα που αντιπροσωπεύουν τις γραμματικές κατηγορίες των λέξεων της πρότασης (O= ουσιαστικό, P= ρήμα, Πρ= Πρόθεση A= άρθρο κ.ο.κ).



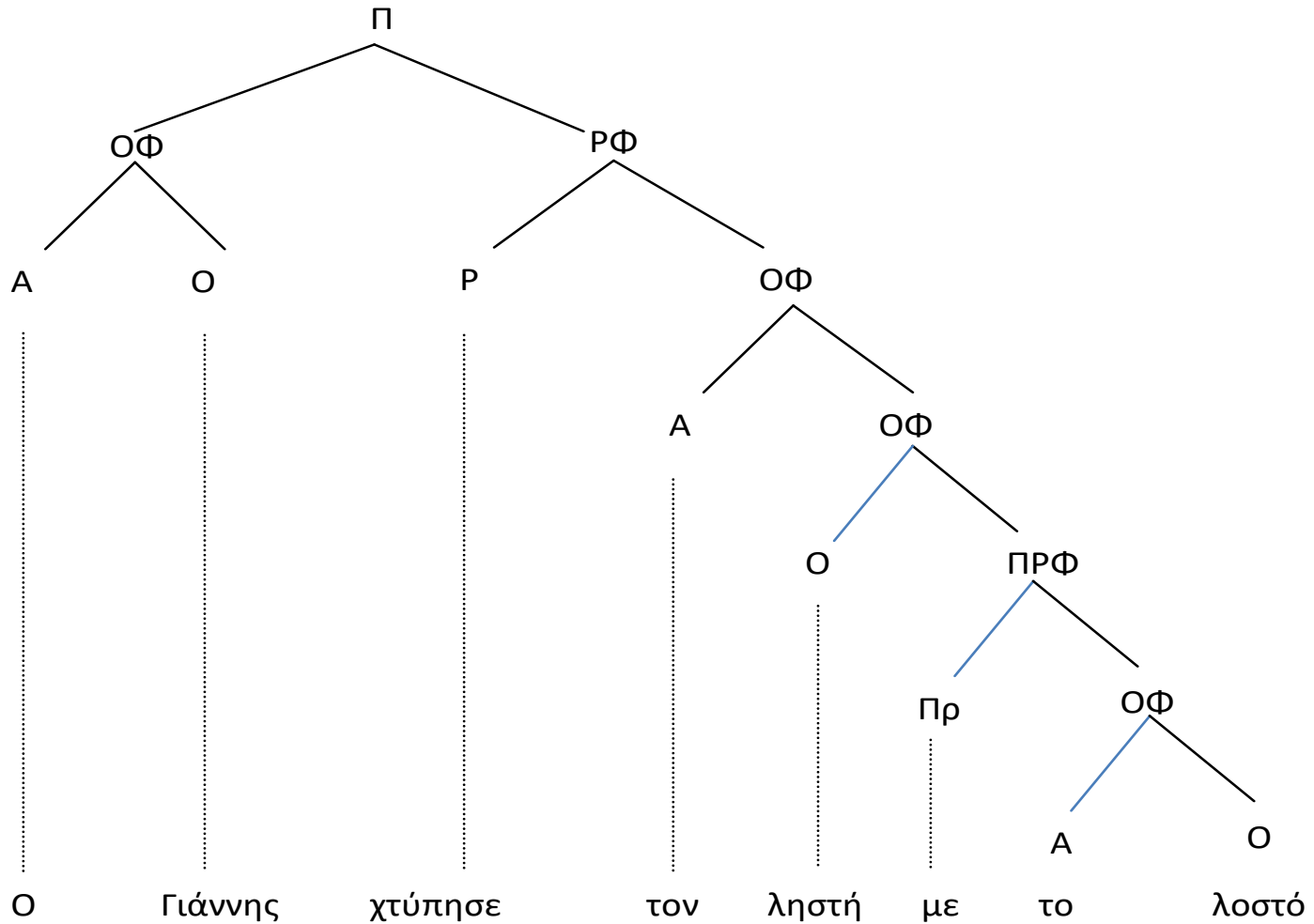
# Γραμματική Φραστικής Δομής

2. Σύμβολα που αντιπροσωπεύουν τις φράσεις της πρότασης (ΟΦ= ονοματική φράση, ΡΦ= ρηματική φράση, ΠρΦ= προθετική φράση κ.ο.κ.)
  3. Σύμβολο που αντιπροσωπεύει την ίδια την πρόταση (Π).
- Κανόνες φραστικής δομής που καθορίζουν τον τρόπο με τον οποίο οι γραμματικές κατηγορίες πρέπει να συνδυαστούν για να σχηματίσουν φράσεις και συνακόλουθα τον τρόπο που οι φράσεις θα συνδυαστούν για να φτιάξουν την πρόταση. Δείγμα κανόνων για την πρόταση που αναλύσαμε παραπάνω είναι οι παρακάτω:
    - (α)  $\Pi \rightarrow \text{ΟΦ} + \text{ΡΦ}$
    - (β)  $\text{ΟΦ} \rightarrow \text{Α} + \text{Ο}$
    - (γ)  $\text{ΡΦ} \rightarrow \text{Ρ} + \text{ΟΦ}$

# Επίλυση σημασιολογικών αμφισημιών στη ΓΦΔ



# Επίλυση σημασιολογικών αμφισημιών στη ΓΦΔ



# Γραμματικές N. Chomsky

- Μια τυπική γραμματική είναι ένα σύστημα παραγωγής συμβολοσειρών  $G$  που αποτελείται από μια τετράδα  $(T, N, P, S)$ .
  - $T$  είναι τα τερματικά σύμβολα (terminals) που αποτελούν το αλφάβητο της γλώσσας και τα οποία μπορούν να είναι η είσοδος ή η έξοδος σε έναν κανόνα παραγωγής και τα οποία δεν μπορούν να αλλάξουν κατά τη χρήση των κανόνων.
  - $N$  είναι τα μη τερματικά σύμβολα (non-terminals) που μπορούν να αλλάξουν. Για παράδειγμα :
    - $A$  μπορεί να γίνει  $B$
    - $A$  μπορεί να γίνει  $X$
  - $P$  είναι οι κανόνες παραγωγής (production rules). Η γραμματική καθορίζεται από αυτούς τους κανόνες και οι οποίοι χρησιμοποιούνται για να αναλύσουν ή να παραγάγουν τις συμβολοσειρές. Κάθε τέτοιος κανόνας παραγωγής έχει την κεφαλή (head) ή την αριστερή πλευρά που είναι η συμβολοσειρά που θα αντικατασταθεί ή θα αναλυθεί και το σώμα ή τη δεξιά πλευρά που είναι η συμβολοσειρά που μπορεί να την αντικαταστήσει.
  - $S$  που είναι το αρχικό σύνολο εκκίνησης (start symbol)



Ευχαριστώ για την προσοχή σας!