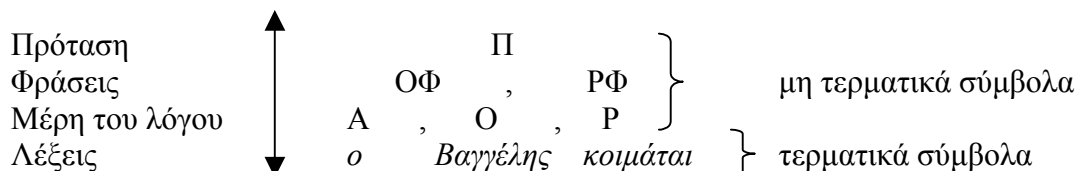


Γραμματικοί κανόνες - Κανόνες μεταγραφής
Ιεραρχία γραμματικών: Γραμματικές Πεπερασμένων Καταστάσεων, Συμφραστικά Ανεξάρτητες Γραμματικές, Συμφραστικά Εξαρτημένες γραμματικές, Γραμματικές Φραστικής Δομής, Μετασχηματιστικές Γραμματικές Παραγωγή συμβολοσειρών – Φραστικοί δείκτες – Δομική αμφισημία



Γραμματικοί κανόνες – Κανόνες μεταγραφής συμβόλων

$A \rightarrow B / \Gamma \underline{\quad} \Delta$,

- όπου: 1. A, B, Γ, Δ είναι σύμβολα ή σειρές συμβόλων (συμβολοσειρές).
 2. A ή B (αλλά όχι A και B) μπορούν να είναι \emptyset (κενά).
 3. Γ, Δ μπορούν να είναι \emptyset .

Εάν Γ ή Δ $\neq \emptyset$, ο κανόνας λέγεται **συμφραστικά εξαρτημένος** (context-sensitive). εάν Γ και Δ = \emptyset , ο κανόνας λέγεται **συμφραστικά ανεξάρτητος** (context-free). Εάν $A \rightarrow X A$, ο κανόνας λέγεται **αναδρομικός (αναδρομικός δεξιά)**. Εάν $A \rightarrow A X$, ο κανόνας λέγεται **αναδρομικός αριστερά**.

- Μια **γραμματική** αποτελείται από (1) ένα αρχικό σύμβολο Π [=Πρόταση], (2α) ένα σύνολο τερματικών συμβόλων, (2β) ένα σύνολο μη τερματικών συμβόλων και (3) ένα σύνολο κανόνων που παράγουν συμβολοσειρές.
- Μια **γλώσσα** είναι ένα (άπειρο) σύνολο τερματικών συμβολοσειρών. Λέμε ότι μια γλώσσα «περιγράφεται» ή «γεννιέται» ή «παράγεται» ή «απαριθμείται» ή «υπολογίζεται» από μια γραμματική.

Είδη γραμματικής

Γραμματική Πεπερασμένων Καταστάσεων:

$\left\{ \begin{array}{l} A \rightarrow \alpha B \text{ (ή } A \rightarrow B \alpha), \\ A \rightarrow \alpha, \end{array} \right.$

όπου τα A, B είναι μη τερματικά (τελικά) σύμβολα και το α είναι τερματικό (τελικό) σύμβολο.

Συμφραστικά Ανεξάρτητη Γραμματική:

$A \rightarrow B$,

όπου το A είναι μη τερματικό σύμβολο και το B είναι μια σειρά τερματικών ή μη τερματικών συμβόλων.

Συμφραστικά Εξαρτημένη Γραμματική:

$\Gamma A \Delta \rightarrow \Gamma B \Delta \text{ (} A \rightarrow B / \Gamma \underline{\quad} \Delta \text{)},$

όπου το A είναι μη τερματικό σύμβολο, τα Γ, Δ είναι σειρές τερματικών ή μη τερματικών συμβόλων και τουλάχιστον ένα από τα Γ, Δ δεν είναι \emptyset . Εναλλακτικά:

$A \rightarrow B$,

όπου A, B είναι δύο οποιεσδήποτε συμβολοσειρές: $|A| \leq |B|$, όπου |X| το μήκος της συμβολοσειράς X.

Γραμματικές Φραστικής Δομής:

$A \rightarrow B$,

όπου A, B είναι δύο οποιεσδήποτε συμβολοσειρές, τερματικών ή μη τερματικών συμβόλων.

Μετασχηματιστικές Γραμματικές:

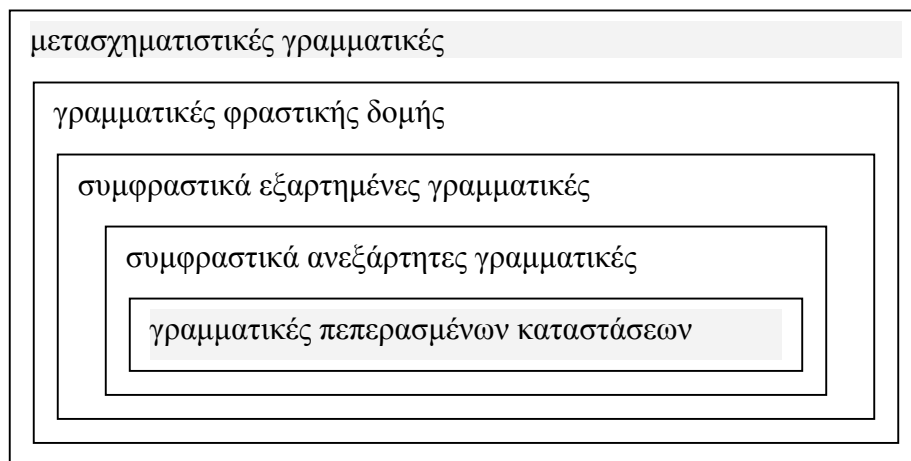
γραμματικές εξαρτημένες ή ανεξάρτητες συμφραστικά οι οποίες, επιπλέον, περιέχουν κανόνες φραστικής δομής

$A \rightarrow B$,

όπου A, B είναι οποιεσδήποτε σειρές τερματικών ή μη τερματικών συμβόλων. Οι κανόνες αυτοί λέγονται **μετασχηματισμοί**.

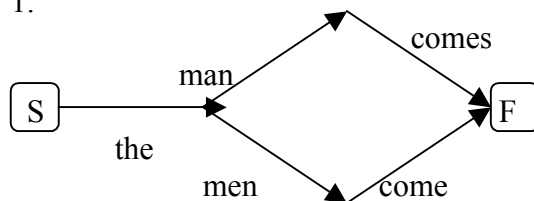
Ιεραρχία των γραμματικών

μετασχηματιστικές γραμματικές \supseteq γραμματικές φραστικής δομής \supseteq συμφραστικά εξαρτημένες γραμματικές \supseteq συμφραστικά ανεξάρτητες γραμματικές \supseteq γραμματικές πεπερασμένων καταστάσεων



Παράδειγματα

1.

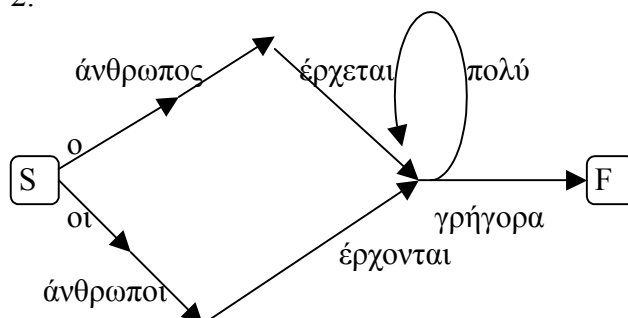


$S \rightarrow \text{the } A, A \rightarrow \text{man } B_1,$
 $A \rightarrow \text{men } B_2,$

$B_1 \rightarrow \text{comes } F,$

$B_2 \rightarrow \text{come } F.$

2.



$S \rightarrow \text{ο } A_1, S \rightarrow \text{οι } A_2,$

$A_1 \rightarrow \text{άνθρωπος } B_1,$

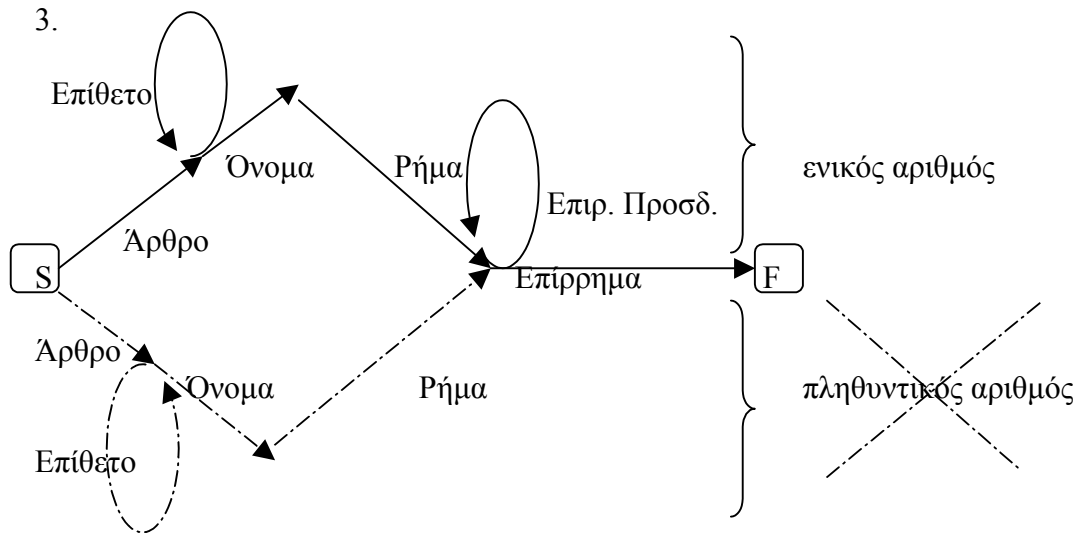
$B_1 \rightarrow \text{έρχεται } \Gamma_1, \Gamma_1 \rightarrow \text{πολύ } \Gamma_1$

$\Gamma_1 \rightarrow \text{γρήγορα } F$

$A_2 \rightarrow \text{άνθρωποι } B_2,$

$B_2 \rightarrow \text{έρχονται } \Gamma_2, \Gamma_2 \rightarrow \text{πολύ } \Gamma_2$

$\Gamma_2 \rightarrow \text{γρήγορα } F.$



- 4.
- $\{αβ, ααββ, αααβββ, \dots\} = \{α^vβ^v: v \geq 1\}$
 Τερματικά σύμβολα: $α, β$
 Μη τερματικά σύμβολα: Π
 Κανόνες:
 1. $\Pi \rightarrow α \Pi β$
 2. $\Pi \rightarrow \emptyset$

Παράδειγμα παραγωγής:

$α \Pi β$
 $α α \Pi β β$
 $α α \emptyset β β$
 $α α β β$

- 5.
- $\{αα, ββ, αββα, βααβ, αααα, ββββ, ααββαα, αββββα, \dots\} = \{ww^R: w = α,β\}$
 Τερματικά σύμβολα: $α, β$
 Μη τερματικά σύμβολα: Π
 Κανόνες:
 1. $\Pi \rightarrow α \Pi α$
 2. $\Pi \rightarrow β \Pi β$
 3. $\Pi \rightarrow \emptyset$

Παράδειγμα παραγωγής:

$α \Pi α$
 $α α \Pi α α$
 $α α β \Pi β α α$
 $α α β α \Pi α β α α$
 $α α β α \emptyset α β α α$
 $α α β α α β α α$

- 6.
- $\{αα, ββ, αβαβ, βαβα, αααα, ββββ, ααβααβ, αββαββ, \dots\} = \{ww: w = α,β\}$
 Τερματικά σύμβολα: $α, β$
 Μη τερματικά σύμβολα: Π, X_1, X_2, X_3, A, B
 Κανόνες:
 1. $\Pi \rightarrow X_1 X_2 X_3$
 2. $X_1 X_2 \rightarrow α X_1 A$
 3. $X_1 X_2 \rightarrow β X_1 B$
 4. $A β \rightarrow β A$
 5. $B α \rightarrow α B$
 6. $A X_3 \rightarrow X_2 α X_3$
 7. $B X_3 \rightarrow X_2 β X_3$
 8. $α X_2 \rightarrow X_2 α$
 9. $β X_2 \rightarrow X_2 β$
 10. $X_1 X_2 \rightarrow \emptyset$
 11. $X_3 \rightarrow \emptyset$

Παράδειγμα παραγωγής:

$X_1 X_2 X_3$
 $α X_1 A X_3$
 $α X_1 X_2 α X_3$
 $α β X_1 B α X_3$
 $α β X_1 α B X_3$
 $α β X_1 α X_2 β X_3$
 $α β X_1 X_2 α β X_3$
 $α β \emptyset α β \emptyset$
 $α β α β$

Θεώρημα 1. Οι φυσικές γλώσσες δεν μπορούν να παραχθούν από γραμματικές πεπερασμένων καταστάσεων.

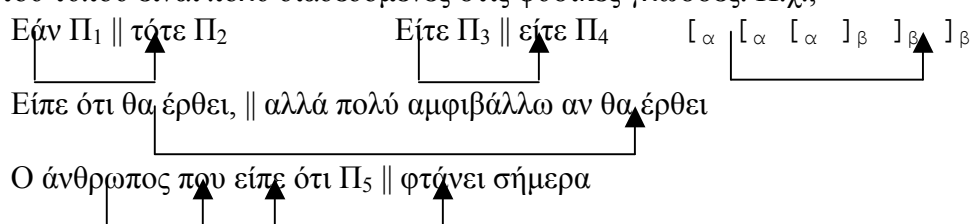
Απόδειξη: Έστω μια γλώσσα που περιέχει τις μονάδες (σύμβολα) a, b . Οι ακόλουθες δομές της γλώσσας αυτής δεν μπορούν να παραχθούν από γραμματικές πεπερασμένων καταστάσεων:

(i) $ab, aabb, aaabbb, \dots : \Gamma = \{a^n b^n : n \geq 1\}$

(ii) $aa, bb, abba, baab, aaaa, bbbb, aabbaa, abbbba, \dots : \Gamma = \{ww^R : w = a, b\}$

(iii) $aa, bb, abab, baab, aaaa, bbbb, aabaaab, abbaab, \dots : \Gamma = \{ww : w = a, b\}$

Η (i) και η (ii) παράγονται από μια συμφραστικά ανεξάρτητη γραμματική. Η (iii) παράγεται από μια γραμματική φραστικής δομής (βλ. παραπάνω). Όμως οι δομές αυτού του τύπου είναι πολύ διαδεδομένες στις φυσικές γλώσσες. Π.χ.,



Ήρθε και ο σκύλος και τρόμαξε το γάτο που έφαγε τον ποντικό που πήρε το φυτίλι μέσα απ' το καντήλι που έφεγγε και κένταγε η κόρη το μαντίλι.

Άρα, οι γραμματικές πεπερασμένων καταστάσεων δεν είναι κατάλληλες για τις φυσικές γλώσσες.

Θεώρημα 2. Οι γραμματικές των φυσικών γλωσσών δεν είναι ισοδύναμες με τις μετασχηματιστικές γραμματικές ή με τις γραμματικές φραστικής δομής.

Απόδειξη (Peters – Ritchie): Οι μετασχηματιστικές γραμματικές και οι γραμματικές φραστικής δομής είναι πολύ ισχυρές για να είναι χρήσιμες στην περιγραφή των φυσικών γλωσσών.

Ασκήσεις

1. Θ. Παυλίδου, *Επίπεδα γλωσσικής ανάλυσης*, σσ. 91-92.
2. Πολλοί ξεχνιούνται στο ταίριασμα των παρενθέσεων. Γράψτε μια γραμματική ισορροπημένων παρενθέσεων, δηλ. μια γραμματική που να παράγει τη γλώσσα: $\{(), (()), \dots, ()(), ()(), \dots, (())(), (())(), \dots, (())((())), \dots, (())((()))(), \dots\}$.
3. Σχεδιάστε φραστικούς δείκτες για τις ακόλουθες δομικά διαφορούμενες φράσεις και προτάσεις:
 - η αγάπη του Θεού
 - η περιγραφή του παιδιού
 - Σκότωσε τον άνθρωπο με το πριόνι
 - πολλοί φοιτητές και ενδιαφερόμενοι καθηγητές
 - Έφτασε νωρίς χτες βράδι.
4. Γράψτε γραμματικές για το ακόλουθο τμήμα της ελληνικής γλώσσας: $\{\text{Ο Γιάννης, που ήρθε χτες, κοιμάται σήμερα} \cdot \text{Η Μαρία, που κοιμόταν χτες, ήρθε σήμερα} \cdot \text{Ο Νίκος, που έγραφε χτες, διαβάζει σήμερα} \cdot \text{Η Γιαννούλα, που έφυγε σήμερα, έγραφε χτες} \cdot \dots\}$
5. Γράψτε τη γραμματική της γλώσσας $\{a^n b^n : n \geq 1\}$. Τι παρατηρείτε σε σχέση με τη γραμματική της γλώσσας $\{a^n b^v : n \geq 1\}$;

$\Gamma = \{\square^v \blacktriangle^v \text{☺}^v : v \geq 1\}$

Τ. σ.: $\square, \blacktriangle, \text{☺}$

Μη τ. σ.: $\Pi, \cdot, -$

Κ.:

1. $\Pi \rightarrow \cdot \blacktriangle \text{☺}$

2. $\cdot \rightarrow \square$

3. $\cdot \rightarrow \square \cdot -$

4. $- \blacktriangle \rightarrow \blacktriangle -$

5. $- \text{☺} \rightarrow \blacktriangle \text{☺} \text{☺}$

