



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΣΥΣΤΗΜΑΤΑ ΚΑΙ ΤΕΧΝΟΛΟΓΙΕΣ ΓΝΩΣΗΣ

Ακ. Έτος 2012-2013

ΑΣΚΗΣΗ 3^η

για τη Σημασιολογία της Γλώσσας των Νηπίων

Δημήτρης Κολοτούρος
03107116

Γιώργος Παρασκευόπουλος
03109030

ΓΕΝΙΚΑ

Σκοπός της εργασίας είναι σημασιολογική ανάλυση της γλώσσας των νηπίων. Ως γλώσσα των νηπίων θεωρούμε τη γλώσσα που αποτελείται από ονόματα και ονοματοποιημένα ρήματα κατά τις προδιαγραφές της εκφώνησης. Θα αναπτύξουμε ένα εργαλείο που θα εξάγει τη σημασία των προτάσεων και θα παράγει μια (ή περισσότερες) προτάσεις αντίστοιχης σημασίας στα ελληνικά, κατά τις προδιαγραφές της εκφώνησης.

Για την υλοποίηση χρησιμοποιήθηκε το εργαλείο GULP (Graph Unification Logic Programming) 4, που προσφέρει υλοποίηση ενοποιητικών γραμματικών σε prolog και η SWI Prolog 6.4.1. Ο κώδικας του GULP 4 διαφέρει απ' την έκδοση 3.1 για λόγους συμβατότητας με την SWI Prolog. Για να χρησιμοποιηθεί ο κώδικας μας σε προηγούμενες εκδόσεις, θα πρέπει να αντικατασταθούν όλες οι εμφανίσεις του συμβόλου (~) με το σύμβολο (:).

Η γενική ιδέα της υλοποίησης είναι η ανάπτυξη μιας ενοποιητικής γραμματικής για την γλώσσα των νηπίων που θα είναι ικανή να εξάγει τη σημασιολογία. *Ως σημασιολογία μιας πρότασης γλώσσας νηπίων ορίζουμε μια ιδιότητα (sem) της οποίας η τιμή είναι μια δομή ιδιοτήτων, που περιλαμβάνει τις ιδιότητες της ενέργειας (pred), του υποκειμένου (sub) και του αντικειμένου (obj) με σαφώς ορισμένες τιμές.* Επίσης, χρειάστηκε η ανάπτυξη μιας ενοποιητικής γραμματικής για την παραγωγή της πρότασης εξόδου, το αναγκαίο υποσύνολο της γραμματικής της ελληνικής γλώσσας.

ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ

Για τα σύμβολα της γραμματικής και τα ονόματα των ιδιοτήτων χρησιμοποιήθηκαν οι παρακάτω συντομογραφίες. Τα σύμβολα που αφορούν τη γραμματική του νηπίου έχουν το επίθεμα `_ni` ενώ τα σύμβολα που αφορούν την ελληνική γραμματική έχουν το επίθεμα `_gr`.

| | |
|---------|--|
| s | sentence (πρόταση) |
| np | noun phrase (ονοματική φράση) |
| vp | verbal phrase (ρηματική φράση) |
| n | noun (όνωμα) |
| v | verb (ρήμα) |
| pronoun | pronoun (μετοχή) |
| art | article (άρθρο) |
| pred | predicate (κατηγορήμα, η σημασιολογία του ρήματος) |
| sub | subject (υποκείμενο) |
| obj | object (αντικείμενο) |
| sem | semantics (σημασιολογία) |
| trans | transitivity (μεταβατικότητα) |
| pers | person (πρόσωπο) |
| case | case (πτώση) |
| nom | nominal (ονομαστική) |
| acc | accusative (αιτιατική) |
| gen | gender (γένος) |
| m | male (αρσενικό) |

f female (θηλυκό)
n neutral (ουδέτερο)

ΕΝΟΠΟΙΗΤΙΚΗ ΓΡΑΜΜΑΤΙΚΗ

Για τη γλώσσα του νηπίου κάναμε τις εξής παραδοχές:

- Τα τερματικά σύμβολα (λέξεις) χωρίζονται σε ονοματικής και ρηματικής έννοιας.
- Τα τερματικά σύμβολα ρηματικής έννοιας έχουν μια ιδιότητα (trans) που έχει τιμή 'TRUE' αν το αντίστοιχο ρήμα έχει μεταβατική έννοια (παίρνει αντικείμενο) ή τιμή 'FALSE' αν δεν έχει μεταβατική έννοια (δεν παίρνει αντικείμενο).
- Με βάση και τις προδιαγραφές της εκφώνησης καταλήξαμε στον παρακάτω διαχωρισμό (κάποιες λέξεις ρηματικής έννοιας όπως η 'μαμ' έχουν trans='FALSE' επειδή δεν υπάρχει λέξη εντός του λεξιλογίου που να μπορεί να παίξει το ρόλο του αντικειμένου):

| Ονοματική Έννοια | Ρηματική Έννοια | |
|------------------|-----------------|---------------|
| | trans:'TRUE' | trans:'FALSE' |
| μαμά | ντα | κακά |
| μπαμπά | | τσίσσα |
| Νινί | | άτα |
| | | μαμ |

Η σύνταξη των προτάσεων ακολουθεί τη μορφή: Όνομα1 Ρήμα Όνομα2, όπου Όνομα1 είναι το υποκείμενο και Όνομα2 το αντικείμενο. Το υποκείμενο μπορεί να παραλείπεται (οπότε εννοείται η έννοια ΕΓΩ (sem:'ME')). Το αντικείμενο υπάρχει μόνο μετά από Ρήμα με μεταβατική έννοια (trans:'TRUE'), ενώ μπορεί επίσης να παραλείπεται, οπότε εννοείται η έννοια ΕΓΩ (sem:'ME').

Καταστρώνουμε, λοιπόν την παρακάτω ενοποιητική γραμματική που σε μια bottom-up προσέγγιση εξάγει την σημασιολογία μιας πρότασης:

$$\left[\begin{array}{c} s_ni \\ sem: \left[\begin{array}{cc} pred: & P \\ sub: & S \\ obj: & O \end{array} \right] \end{array} \right] \rightarrow \left[\begin{array}{c} n_ni \\ sem: S \end{array} \right] \quad \left[\begin{array}{c} vp_ni \\ sem: \left[\begin{array}{cc} pred: & P \\ obj: & O \end{array} \right] \end{array} \right]$$

$$\left[\begin{array}{c} s_ni \\ sem: \left[\begin{array}{cc} pred: & P \\ sub: & S \\ obj: & 'ME' \end{array} \right] \end{array} \right] \rightarrow \left[\begin{array}{c} vp_ni \\ sem: \left[\begin{array}{cc} pred: & P \\ obj: & O \end{array} \right] \end{array} \right]$$

$$\left[\begin{array}{c} \text{vp_ni} \\ \text{sem:} \left[\begin{array}{cc} \text{pred:} & P \\ \text{obj:} & O \end{array} \right] \end{array} \right] \rightarrow \left[\begin{array}{c} \text{v_ni} \\ \text{sem:} \quad P \\ \text{trans} \quad 'TRUE' \end{array} \right] \quad \left[\begin{array}{c} \text{n_ni} \\ \text{sem:} \quad O \end{array} \right]$$

$$\left[\begin{array}{c} \text{vp_ni} \\ \text{sem:} \left[\begin{array}{cc} \text{pred:} & P \\ \text{obj:} & 'ME' \end{array} \right] \end{array} \right] \rightarrow \left[\begin{array}{c} \text{v_ni} \\ \text{sem:} \quad P \\ \text{trans} \quad 'TRUE' \end{array} \right]$$

$$\left[\begin{array}{c} \text{vp_ni} \\ \text{sem:} \left[\begin{array}{cc} \text{pred:} & P \\ \text{obj:} & 'NONE' \end{array} \right] \end{array} \right] \rightarrow \left[\begin{array}{c} \text{v_ni} \\ \text{sem:} \quad P \\ \text{trans} \quad 'FALSE' \end{array} \right]$$

$$\left[\begin{array}{c} \text{n_ni} \\ \text{sem:} \quad 'MPAMPAS' \end{array} \right] \rightarrow \text{μπαμπά}$$

$$\left[\begin{array}{c} \text{n_ni} \\ \text{sem:} \quad 'MAMA' \end{array} \right] \rightarrow \text{μαμά}$$

$$\left[\begin{array}{c} \text{n_ni} \\ \text{sem:} \quad 'ME' \end{array} \right] \rightarrow \text{νινί}$$

$$\left[\begin{array}{c} \text{v_ni} \\ \text{sem:} \quad 'KAKA' \\ \text{trans:} \quad 'FALSE' \end{array} \right] \rightarrow \text{κακά}$$

$$\left[\begin{array}{c} \text{v_ni} \\ \text{sem:} \quad 'TSISA' \\ \text{trans:} \quad 'FALSE' \end{array} \right] \rightarrow \text{τσίσσα}$$

$$\left[\begin{array}{c} \text{v_ni} \\ \text{sem:} \quad 'NTA' \\ \text{trans:} \quad 'TRUE' \end{array} \right] \rightarrow \text{ντα}$$

$$\left[\begin{array}{c} \text{v_ni} \\ \text{sem:} \quad 'ATA' \\ \text{trans:} \quad 'FALSE' \end{array} \right] \rightarrow \text{άτα}$$

$$\left[\begin{array}{c} \text{v_ni} \\ \text{sem:} \quad 'MAM' \\ \text{trans:} \quad 'FALSE' \end{array} \right] \rightarrow \text{μाम}$$

Για την παραγωγή της σημασίας της κάθε πρότασης ήταν αναγκαίο να καταστρωθεί μια ενοποιητική γραμματική με το υποσύνολο των γραμματικών κανόνων της ελληνικής γλώσσας που είναι απαραίτητοι για την παραγωγή των προτάσεων που έχουν την ίδια σημασιολογία με τις προτάσεις που μπορούν να παραχθούν από την γλώσσα των νηπίων.

Καταστρώσαμε λοιπόν την παρακάτω γραμματική:

$$\begin{bmatrix} \text{s_gr} \\ \text{sem:} \begin{bmatrix} \text{pred: } P \\ \text{sub: } S \\ \text{obj: } O \end{bmatrix} \end{bmatrix} \rightarrow \begin{bmatrix} \text{np_gr} \\ \text{case: } \text{nom} \\ \text{sem: } S \end{bmatrix} \begin{bmatrix} \text{vp_gr} \\ \text{pers: } '3' \\ \text{sem:} \begin{bmatrix} \text{pred: } P \\ \text{obj: } O \end{bmatrix} \end{bmatrix}$$

$$\begin{bmatrix} \text{s_gr} \\ \text{sem:} \begin{bmatrix} \text{pred: } P \\ \text{sub: } 'ME' \\ \text{obj: } O \end{bmatrix} \end{bmatrix} \rightarrow \begin{bmatrix} \text{vp_gr} \\ \text{pers: } '1' \\ \text{sem:} \begin{bmatrix} \text{pred: } P \\ \text{obj: } O \end{bmatrix} \end{bmatrix}$$

$$\begin{bmatrix} \text{np_gr} \\ \text{case: } C \\ \text{sem: } S \end{bmatrix} \rightarrow \begin{bmatrix} \text{art_gr} \\ \text{gen: } G \\ \text{case: } C \end{bmatrix} \begin{bmatrix} \text{n_gr} \\ \text{gen: } G \\ \text{case: } C \\ \text{sem: } S \end{bmatrix}$$

$$\begin{bmatrix} \text{vp_gr} \\ \text{pers: } PERS \\ \text{sem:} \begin{bmatrix} \text{pred: } P \\ \text{obj: } O \end{bmatrix} \end{bmatrix} \rightarrow \begin{bmatrix} \text{v_gr} \\ \text{pers: } PERS \\ \text{sem: } P \end{bmatrix} \begin{bmatrix} \text{np_gr} \\ \text{case: } acc \\ \text{sem: } O \end{bmatrix}$$

$$\begin{bmatrix} \text{vp_gr} \\ \text{pers: } PERS \\ \text{sem:} \begin{bmatrix} \text{pred: } P \\ \text{obj: } O \end{bmatrix} \end{bmatrix} \rightarrow \begin{bmatrix} \text{pronoun_gr} \\ \text{case: } acc \\ \text{sem: } O \end{bmatrix} \begin{bmatrix} \text{v_gr} \\ \text{pers: } PERS \\ \text{sem: } P \end{bmatrix}$$

$$\begin{bmatrix} \text{vp_gr} \\ \text{pers: } PERS \\ \text{sem:} \begin{bmatrix} \text{pred: } P \\ \text{obj: } 'NONE' \end{bmatrix} \end{bmatrix} \rightarrow \begin{bmatrix} \text{v_gr} \\ \text{pers: } PERS \\ \text{sem: } P \end{bmatrix}$$

$$\begin{bmatrix} \text{n_gr} \\ \text{gen: } m \\ \text{case: } nom \\ \text{sem: } 'MPAMPAS' \end{bmatrix} \rightarrow \text{μπαμπάς}$$

| | | |
|---|---|--------|
| $\begin{bmatrix} \text{gen:} & n_gr & m \\ \text{case:} & acc \\ \text{sem} & 'MPAMPAS' \end{bmatrix}$ | → | μπαμπά |
|---|---|--------|

| | | |
|--|---|------|
| $\begin{bmatrix} \text{gen:} & n_gr & f \\ \text{sem} & 'MAMA' \end{bmatrix}$ | → | μαμά |
|--|---|------|

| | | |
|---|---|----|
| $\begin{bmatrix} \text{pronoun_gr} \\ \text{case:} & acc \\ \text{sem} & 'ME' \end{bmatrix}$ | → | με |
|---|---|----|

| | | |
|---|---|-----------|
| $\begin{bmatrix} \text{v_gr} \\ \text{pers:} & '1' \\ \text{sem} & 'KAKA' \end{bmatrix}$ | → | κάνω κακά |
|---|---|-----------|

| | | |
|---|---|-------------------|
| $\begin{bmatrix} \text{v_gr} \\ \text{pers:} & '1' \\ \text{sem} & 'KAKA' \end{bmatrix}$ | → | θέλω να κάνω κακά |
|---|---|-------------------|

| | | |
|---|---|------------|
| $\begin{bmatrix} \text{v_gr} \\ \text{pers:} & '3' \\ \text{sem} & 'KAKA' \end{bmatrix}$ | → | κάνει κακά |
|---|---|------------|

| | | |
|--|---|------------|
| $\begin{bmatrix} \text{v_gr} \\ \text{pers:} & '1' \\ \text{sem} & 'TSISA' \end{bmatrix}$ | → | κάνω τσίσα |
|--|---|------------|

| | | |
|--|---|--------------------|
| $\begin{bmatrix} \text{v_gr} \\ \text{pers:} & '1' \\ \text{sem} & 'TSISA' \end{bmatrix}$ | → | θέλω να κάνω τσίσα |
|--|---|--------------------|

| | | |
|--|---|-------------|
| $\begin{bmatrix} \text{v_gr} \\ \text{pers:} & '3' \\ \text{sem} & 'TSISA' \end{bmatrix}$ | → | κάνει τσίσα |
|--|---|-------------|

| | | |
|--|---|-------|
| $\begin{bmatrix} \text{v_gr} \\ \text{pers:} & '1' \\ \text{sem} & 'NTA' \end{bmatrix}$ | → | δέρνω |
|--|---|-------|

v_gr
[pers: '3'
sem 'NTA']

→ δέρνει

v_gr
[pers: '1'
sem 'ATA']

→ πήγα βόλτα

v_gr
[pers: '1'
sem 'ATA']

→ θέλω βόλτα

v_gr
[pers: '3'
sem 'ATA']

→ πήγε βόλτα

v_gr
[pers: '1'
sem 'MAM']

→ τρώω

v_gr
[pers: '1'
sem 'MAM']

→ θέλω να φάω

v_gr
[pers: '3'
sem 'MAM']

→ τρώει

art_gr
[gem: m
case: nom]

→ ο

art_gr
[gem: m
case: acc]

→ τον

art_gr
[gem: f
case: nom]

→ η

| | | |
|---|---|-----|
| art_gr | → | την |
| $\begin{bmatrix} gem: & f \\ case: & acc \end{bmatrix}$ | | |

| | | |
|--|---|----|
| art_gr | → | το |
| $\begin{bmatrix} gem: & n \end{bmatrix}$ | | |

ΥΛΟΠΟΙΗΣΗ

Η υλοποίησή μας περιέχει την γραμματική που καταστρώθηκε παραπάνω γραμμένη στον φορμαλισμό του GULP 4. Οι ελληνικές λέξεις αναπαρίστανται σε μορφή greeklish για λόγους συμβατότητας.

Επίσης, η υλοποίηση περιλαμβάνει το κατηγορήμα **prolog ni2gr/2** που δέχεται ως όρισμα μια πρόταση στη γλώσσα των νηπίων και βγάζει ως αποτέλεσμα την συντακτική δομή (τη σημασιολογία της πρότασης) καθώς και τη σημασία της. Κάποιες προτάσεις έχουν πολλαπλές σημασίες, που παρουσιάζονται ως διαφορετικές ενοποιήσεις της prolog (κάθε νέα σημασία εμφανίζεται αφού πληκτρολογήσουμε τον χαρακτήρα ‘;’ στην prolog).

Παραδείγματα χρήσης:

```
?- ni2gr([mama,kaka]).
[h,mama,kanei,kaka]
pred ~ KAKA
sub ~ MAMA
obj ~ NONE
true ;
false.
```

```
?- ni2gr([mpampa,nta]).
[o,mpampas,me,dernei]
pred ~ NTA
sub ~ MPAMPAS
obj ~ ME
true ;
false.
```

```
?- ni2gr([nini,ata]).
[piga,volta]
pred ~ ATA
sub ~ ME
obj ~ NONE
true ;
[thelw,volta]
pred ~ ATA
sub ~ ME
obj ~ NONE
true ;
false.
```



```

?- ni2gr([tsisa]).
[kanw,tsisa]
pred ~ TSISA
sub ~ ME
obj ~ NONE
true ;
[thelw,na,kanw,tsisa]
pred ~ TSISA
sub ~ ME
obj ~ NONE
true ;
false.

?- ni2gr([mam]).
[trww]
pred ~ MAM
sub ~ ME
obj ~ NONE
true ;
[thelw,na,faw]
pred ~ MAM
sub ~ ME
obj ~ NONE
true ;
false.

?- ni2gr([mpampa,nta,nini]).
[o,mpampas,me,dernei]
pred ~ NTA
sub ~ MPAMPAS
obj ~ ME
true ;
false.

```

Για τη χρήση του κώδικα που ακολουθεί, θα πρέπει πρώτα να έχουμε εισάγει το GULP στην Prolog (το αρχείο `gulp4swi.pl` στην περίπτωσή μας).

nipio.glp:

```

s_ni(sem~(pred~P..sub~S..obj~O)) --> n_ni(sem~S), vp_ni(sem~(pred~P..obj~O)).
s_ni(sem~(pred~P..sub~'ME'..obj~O)) --> vp_ni(sem~(pred~P..obj~O)).

vp_ni(sem~(pred~P..obj~O)) --> v_ni(sem~P..trans~'TRUE'), n_ni(sem~O).
vp_ni(sem~(pred~P..obj~'ME')) --> v_ni(sem~P..trans~'TRUE').
vp_ni(sem~(pred~P..obj~'NONE')) --> v_ni(sem~P..trans~'FALSE').

n_ni(sem~'MPAMPAS') --> [mpampa].
n_ni(sem~'MAMA') --> [mama].
n_ni(sem~'ME') --> [nini].

v_ni(sem~'KAKA'..trans~'FALSE') --> [kaka].
v_ni(sem~'TSISA'..trans~'FALSE') --> [tsisa].
v_ni(sem~'NTA'..trans~'TRUE') --> [nta].
v_ni(sem~'ATA'..trans~'FALSE') --> [ata].
v_ni(sem~'MAM'..trans~'FALSE') --> [mam].

```

```

s_gr(sem~(pred~P..sub~S..obj~O)) --> np_gr(case~nom..sem~S),
vp_gr(pers~'3'..sem~(pred~P..obj~O)).
s_gr(sem~(pred~P..sub~'ME'..obj~O)) --> vp_gr(pers~'1'..sem~(pred~P..obj~O)).

np_gr(case~C..sem~S) --> art_gr(gen~G..case~C), n_gr(gen~G..case~C..sem~S).

vp_gr(pers~PERS..sem~(pred~P..obj~O)) --> v_gr(pers~PERS..sem~P),
np_gr(case~acc..sem~O).
vp_gr(pers~PERS..sem~(pred~P..obj~O)) --> pronoun_gr(case~acc..sem~O),
v_gr(pers~PERS..sem~P).
vp_gr(pers~PERS..sem~(pred~P..obj~'NONE')) --> v_gr(pers~PERS..sem~P).

n_gr(gen~m..case~nom..sem~'MPAMPAS') --> [mpampas].
n_gr(gen~m..case~acc..sem~'MPAMPAS') --> [mpampa].
n_gr(gen~f..sem~'MAMA') --> [mama].

pronoun_gr(case~acc..sem~'ME') --> [me].

v_gr(pers~'1'..sem~'KAKA') --> [kanw,kaka].
v_gr(pers~'1'..sem~'KAKA') --> [thelw,na,kanw,kaka].
v_gr(pers~'3'..sem~'KAKA') --> [kanei,kaka].

v_gr(pers~'1'..sem~'TSISA') --> [kanw,tsisa].
v_gr(pers~'1'..sem~'TSISA') --> [thelw,na,kanw,tsisa].
v_gr(pers~'3'..sem~'TSISA') --> [kanei,tsisa].

v_gr(pers~'1'..sem~'NTA') --> [dernw].
v_gr(pers~'3'..sem~'NTA') --> [dernei].

v_gr(pers~'1'..sem~'ATA') --> [piga,volta].
v_gr(pers~'1'..sem~'ATA') --> [thelw,volta].
v_gr(pers~'3'..sem~'ATA') --> [pige,volta].

v_gr(pers~'1'..sem~'MAM') --> [trww].
v_gr(pers~'1'..sem~'MAM') --> [thelw,na,faw].
v_gr(pers~'3'..sem~'MAM') --> [trwei].

art_gr(gen~m..case~nom) --> [o].
art_gr(gen~m..case~acc) --> [ton].
art_gr(gen~f..case~nom) --> [h].
art_gr(gen~f..case~acc) --> [thn].
art_gr(gen~n) --> [to].

ni2gr(String) :- phrase(s_ni(sem~SEM), String), phrase(s_gr(sem~SEM), Output),
writeln([Output]), display_feature_structure(SEM).

```