DOROTA LIPOWSKA
Zakład Logiki Stosowanej, Instytut Językoznawstwa UAM

Wprowadzenie do HPSG
(Head-driven Phrase Structure Grammar)

Adam Przepiórkowski, Anna Kupść, Małgorzata Marcinek i Agnieszka Mykowiecka
„*Formalny opis języka polskiego. Teoria i implementacja*”
Seria: Problemy współczesnej nauki.
Teoria i zastosowania.
Inżynieria lingwistyczna.
Akademicka Oficyna Wydawnicza EXIT, Warszawa 2002.

- język formalny i teoria lingwistyczna

- mechanizmy: hierarchie typów
wielodziedziczenie
wielopoziomowe struktury

-wieloaspektowy opis języka: składnia
semantyka
pragmatyka
fonologia

- teorie generatywne

- Generalized Phrase Structure Grammar

- ograniczenia („*constraint-based*”)

- Lexical Functional Grammar

- **gramatyczne** – spełniające wszystkie ograniczenia

- Carl Pollard (1984) „*Generalized Phrase Structure Grammars, head grammars and natural languages*”

- angielski
- niemiecki
- holenderski
- francuski
- włoski
- hiszpański
- portugalski
- serbsko-chorwacki
- bułgarski
- czeski
- słoweński
- polski

- walijski
- koreański
- japoński
- turecki
- grecki
- hebrajski
- warlpiri
- amerykański język migowy

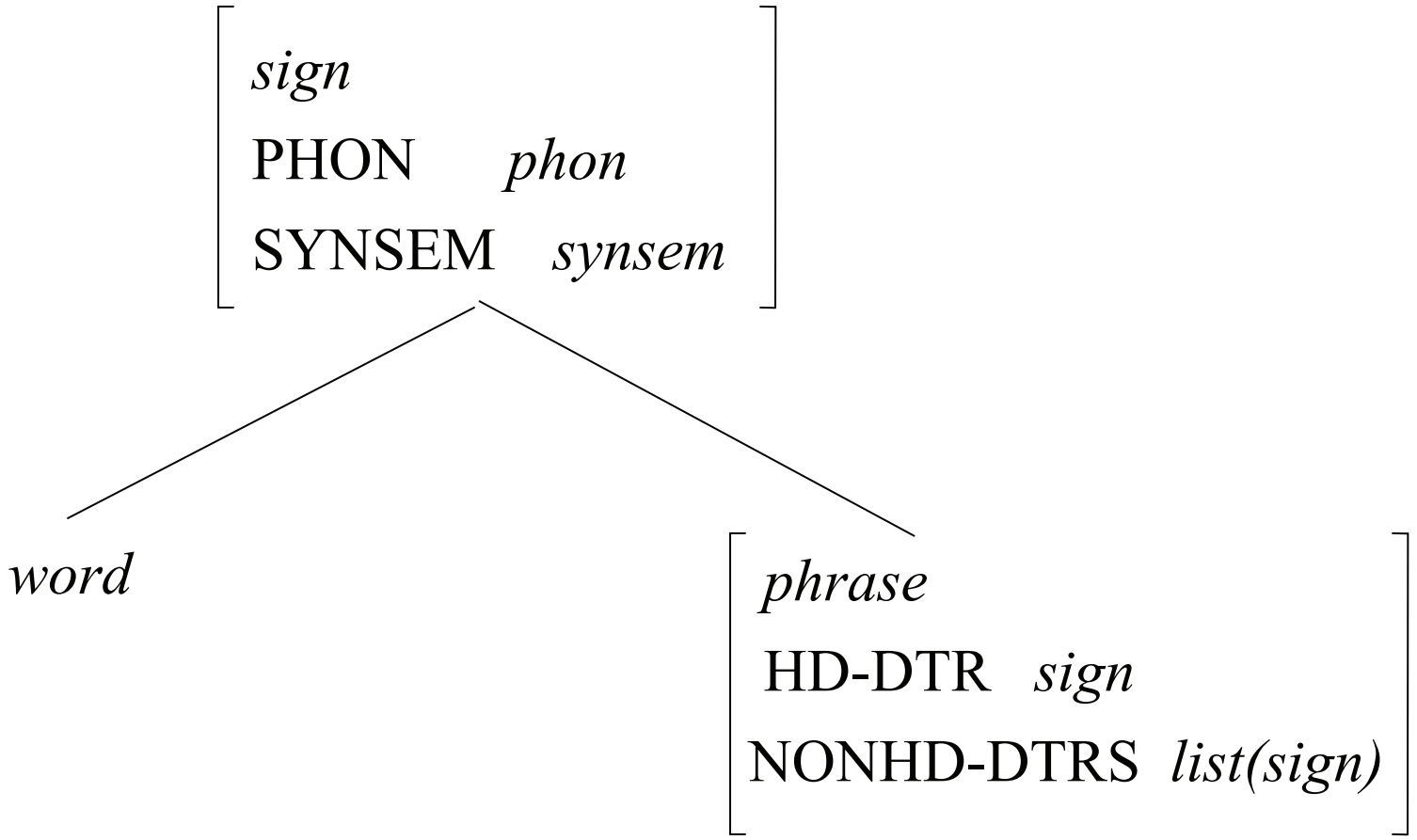
Formalna teoria HPSG:

- sygnatura
- teoria właściwa

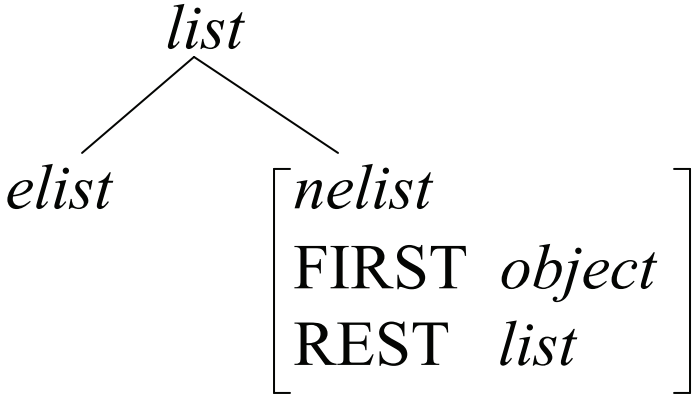
Teoria właściwa – zbiór ograniczeń

Sygnatura:

- częściowy porządek na zbiorze typów
- przyporządkowanie typom atrybutów
- relacje



dziedziczenie atrybutów



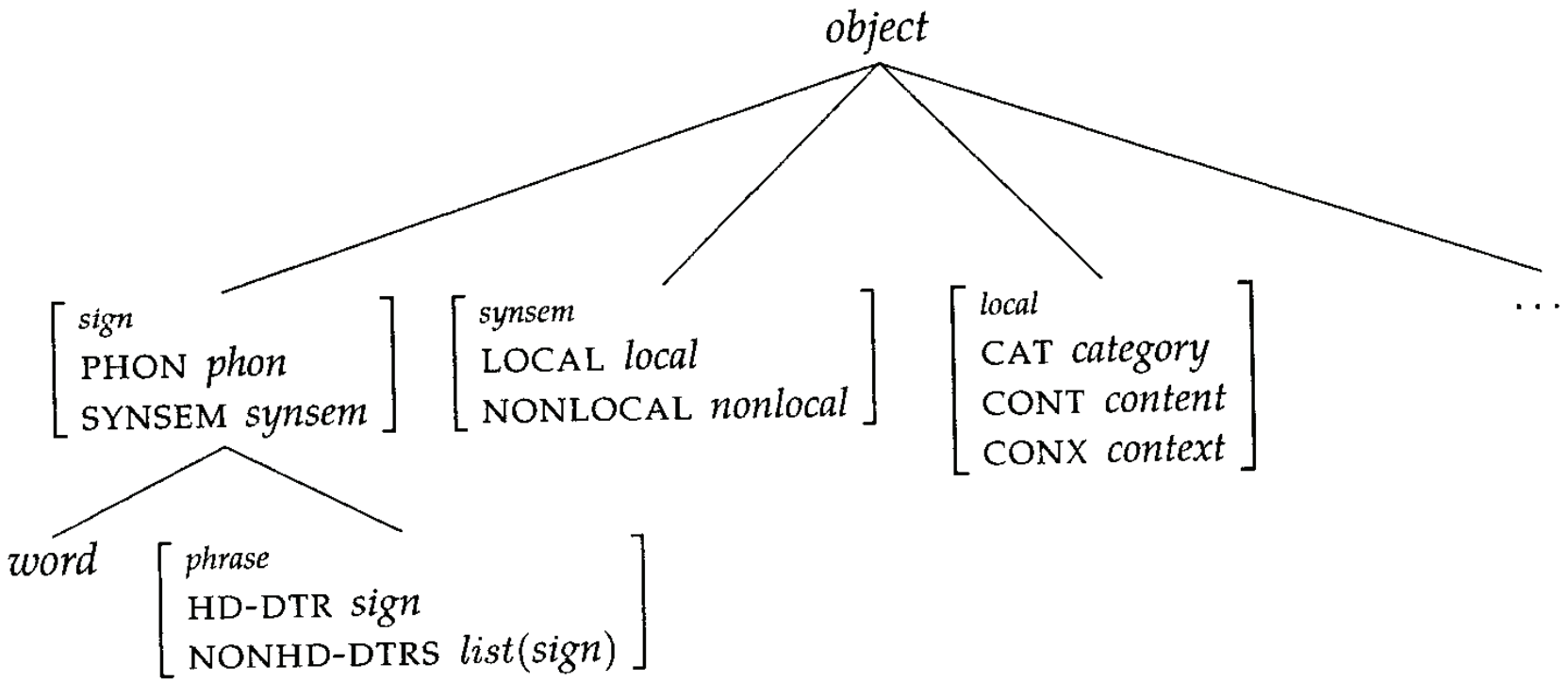
$< > = elist$

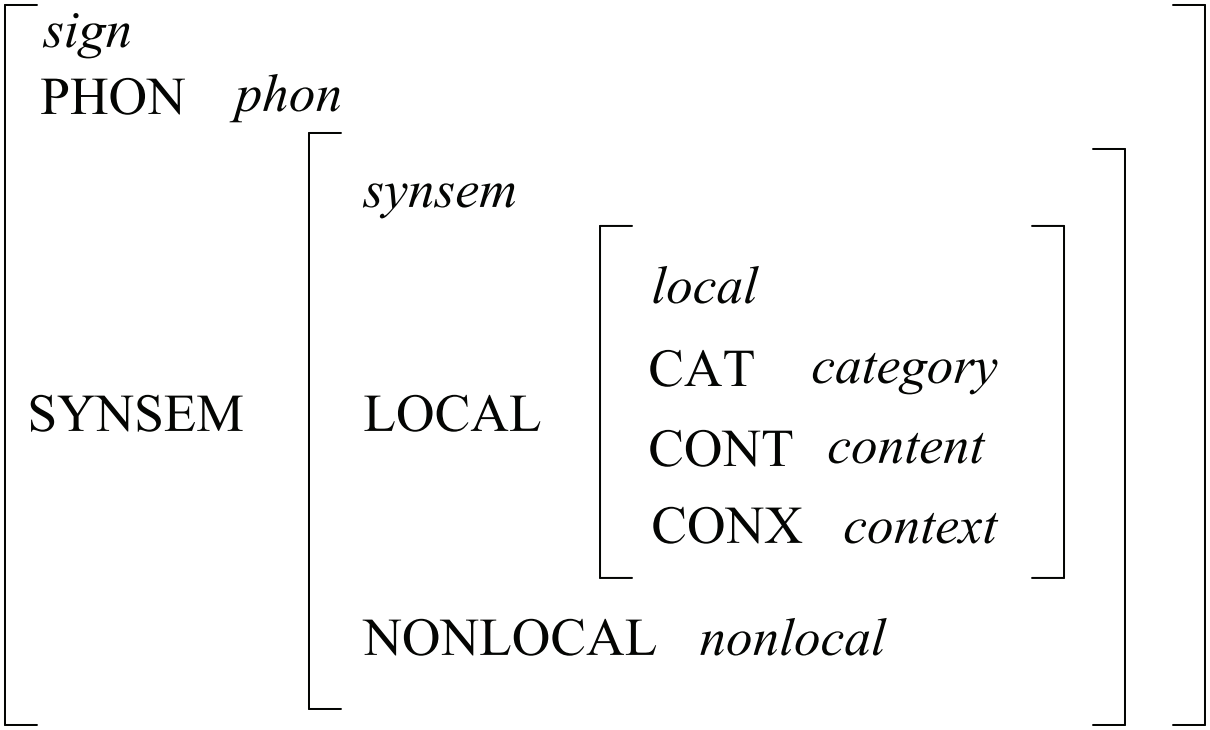
$L = < a_1, a_2, a_3 > = < a_1 \mid R >$

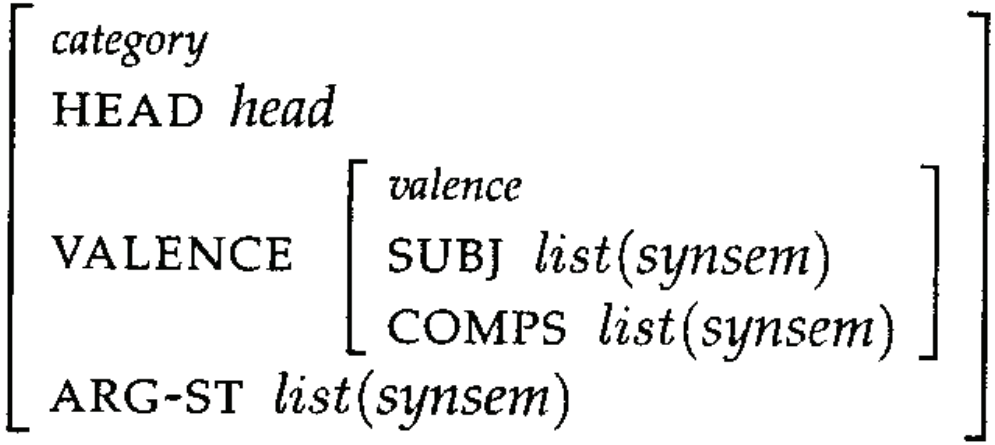
$R = < a_2, a_3 > = < a_2 \mid P >$

$P = < a_3 > = < a_3 \mid S >$

$S = < >$


object
sign
PHON phon
SYNSEM synsem
synsem
LOCAL local
NONLOCAL nonlocal
local
CAT category
CONT content
CONX context
...
word
phrase
HD-DTR sign
NONHD-DTRS list(sign)


sign
PHON phon
SYNSEM
synsem
LOCAL
local
CAT category
CONT content
CONX context
NONLOCAL nonlocal


category
HEAD head
VALENCE
valence
SUBJ list(synsem)
COMPS list(synsem)
ARG-ST list(synsem)

Zasada elementu głównego

$$\textit{phrase} \rightarrow \begin{bmatrix} \text{SYNSEM|LOCAL|CAT|HEAD } \boxed{1} \\ \text{HD-DTR|SYNSEM|LOCAL|CAT|HEAD } \boxed{1} \end{bmatrix}$$

$$\begin{bmatrix} \textit{phrase} \\ \text{SYNSEM} \begin{bmatrix} \textit{synsem} \\ \text{LOCAL} \begin{bmatrix} \textit{local} \\ \text{CAT} \begin{bmatrix} \textit{category} \\ \text{HEAD } \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{HD-DTR} \begin{bmatrix} \textit{sign} \\ \text{SYNSEM} \begin{bmatrix} \textit{synsem} \\ \text{LOCAL} \begin{bmatrix} \textit{local} \\ \text{CAT} \begin{bmatrix} \textit{category} \\ \text{HEAD } \boxed{1} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Zasada Struktury Argumentów

$$word \rightarrow \left[\text{SYNSEM|LOCAL|CAT} \left[\begin{array}{l} category \\ \text{VALENCE} \left[\begin{array}{l} \text{SUBJ } \boxed{1} \\ \text{COMPS } \boxed{2} \end{array} \right] \\ \text{ARG-ST } \boxed{3} \end{array} \right] \right]$$

$$\wedge \ \text{append}(\boxed{1}, \boxed{2}, \boxed{3})$$

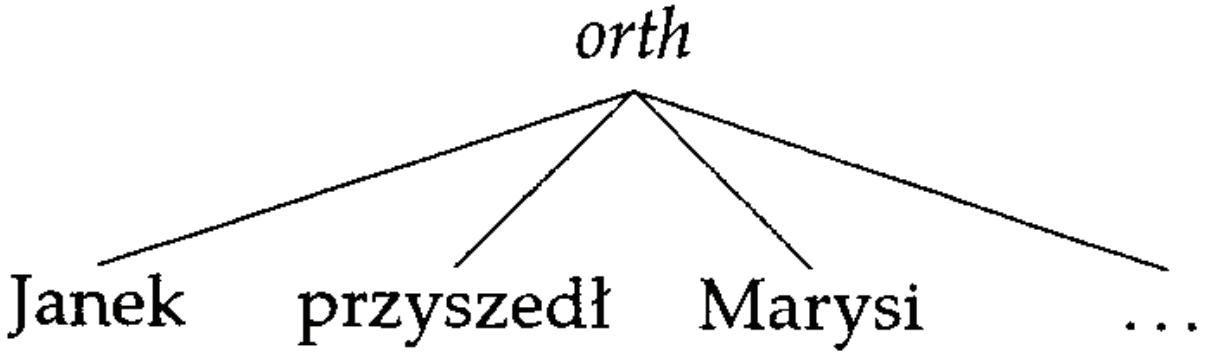
$$\text{append}(\langle\rangle, \boxed{1}, \boxed{1}).$$

$$\text{append}(\langle \boxed{1} | \boxed{2} \rangle, \boxed{3}, \langle \boxed{1} | \boxed{4} \rangle) \overset{\forall}{\Leftarrow} \text{append}(\boxed{2}, \boxed{3}, \boxed{4}).$$

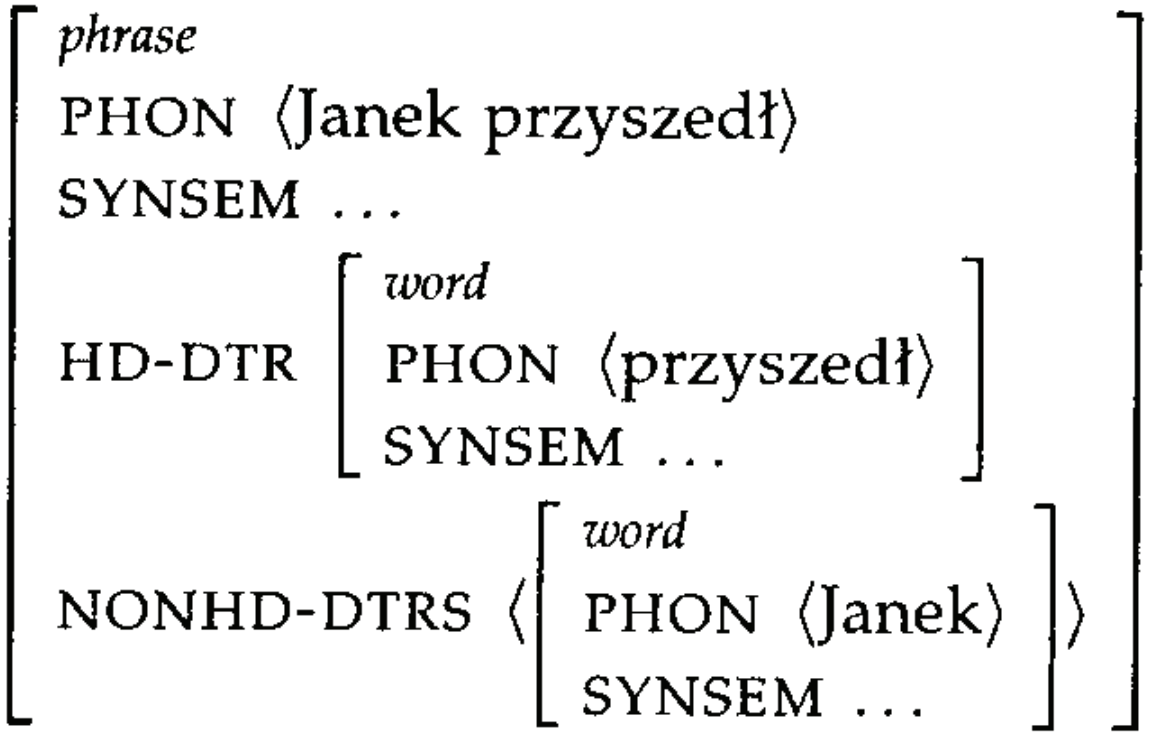
Zasada co najwyżej jednego podmiotu

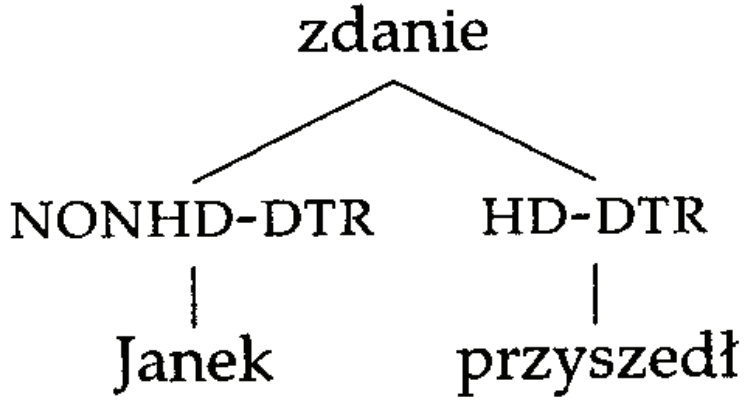
$$\left[\begin{array}{l} \textit{valence} \\ \text{SUBJ } \boxed{1} \end{array}\right] \longrightarrow \left(\boxed{1} = \langle\rangle \vee \boxed{1} = \langle \boxed{0} \rangle\right)$$

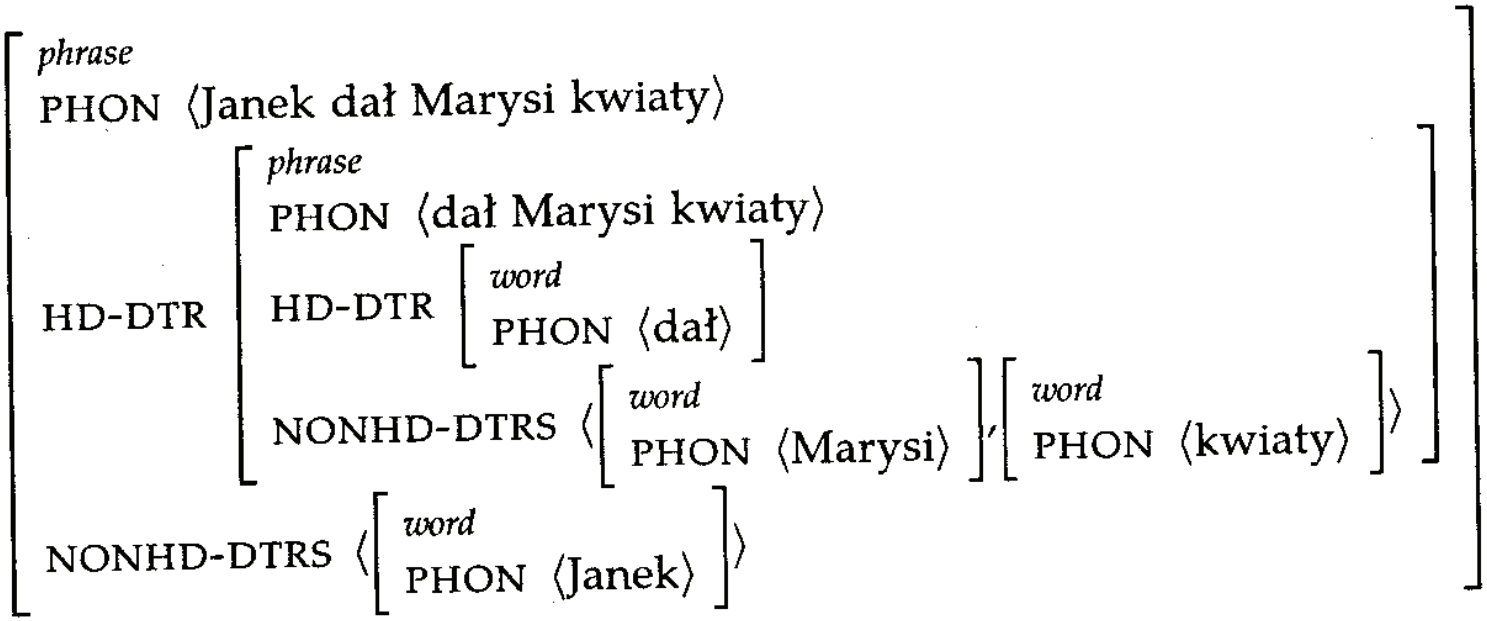
$$\begin{bmatrix} \textit{sign} \\ \text{PHON } \textit{list}(\textit{orth}) \\ \text{SYNSEM } \textit{synsem} \end{bmatrix}$$

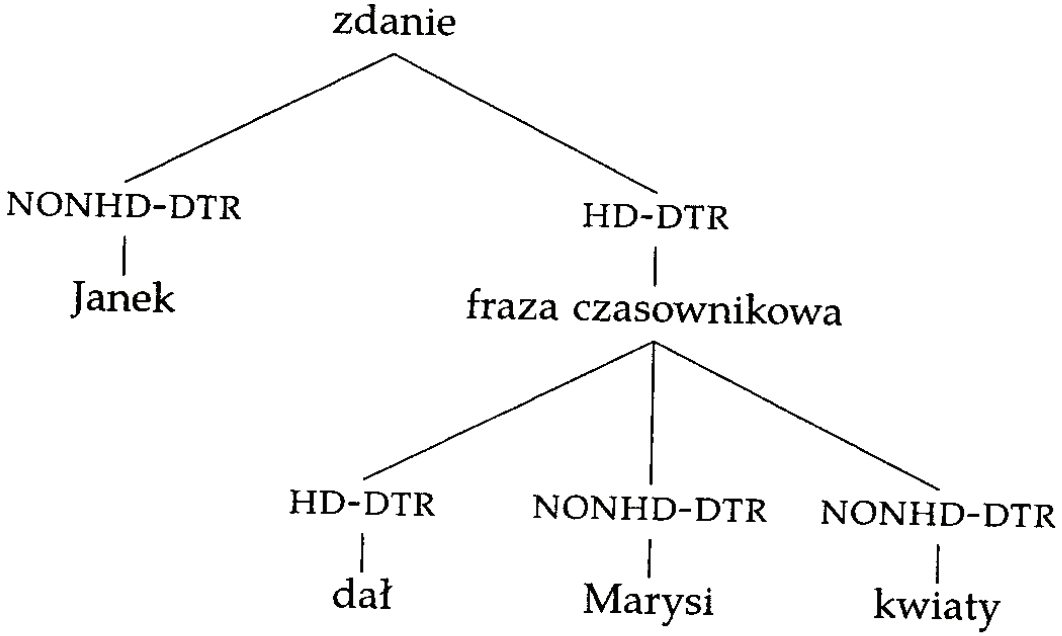


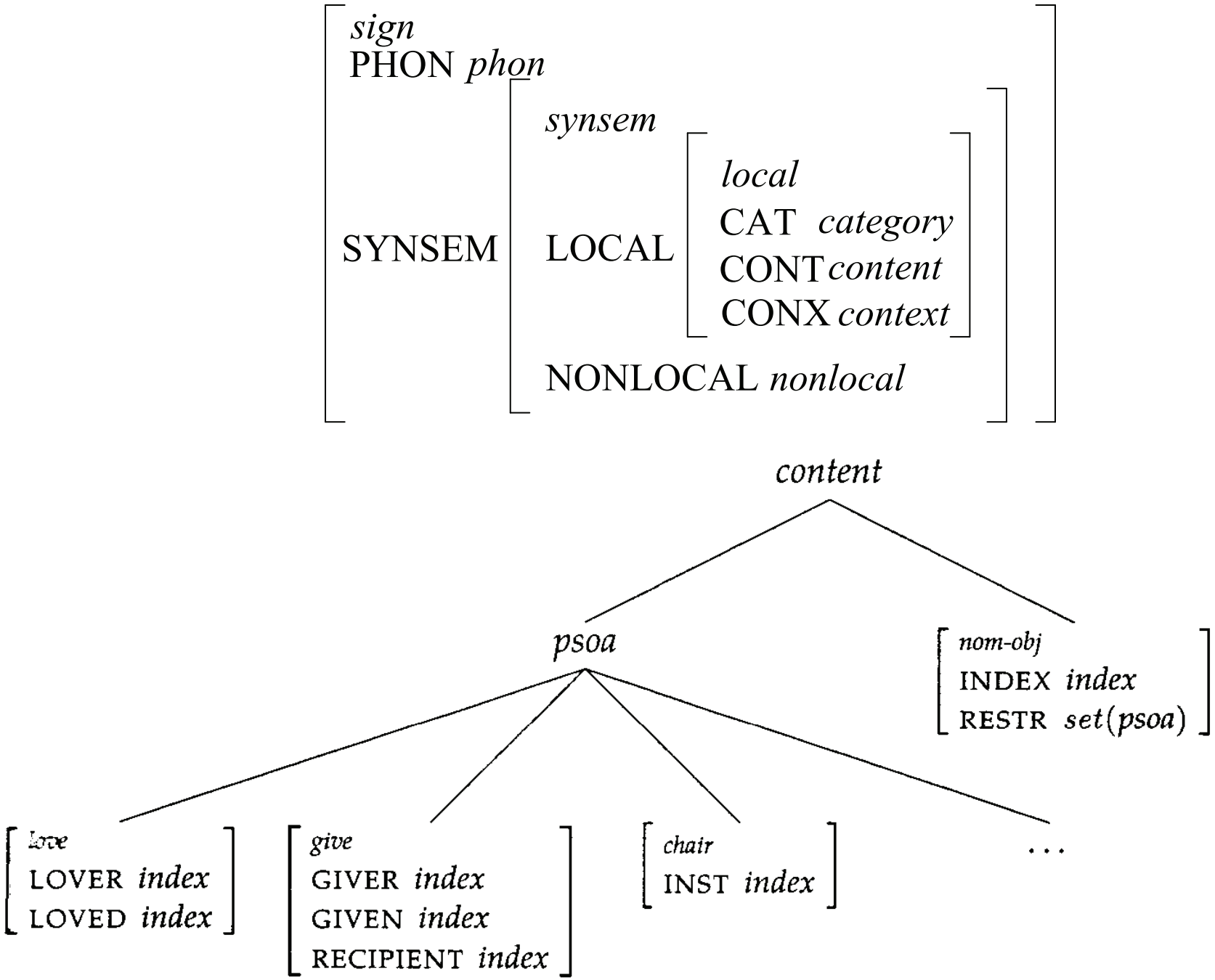
$$\begin{bmatrix} \textit{word} \\ \text{PHON } \langle \text{Janek} \rangle \\ \text{SYNSEM } \ldots \end{bmatrix}$$


phrase
PHON ⟨Janek przyszedł⟩
SYNSEM ...
HD-DTR
word
PHON ⟨przyszedł⟩
SYNSEM ...
NONHD-DTRS ⟨
word
PHON ⟨Janek⟩
SYNSEM ...
⟩


zdanie
NONHD-DTR
HD-DTR
Janek
przyszedł


phrase
PHON ⟨Janek dał Marysi kwiaty⟩
HD-DTR
phrase
PHON ⟨dał Marysi kwiaty⟩
HD-DTR
word
PHON ⟨dał⟩
NONHD-DTRS
word
PHON ⟨Marysi⟩
word
PHON ⟨kwiaty⟩
NONHD-DTRS
word
PHON ⟨Janek⟩


zdanie
NONHD-DTR
HD-DTR
Janek
fraza czasownikowa
HD-DTR
NONHD-DTR
NONHD-DTR
dał
Marysi
kwiaty

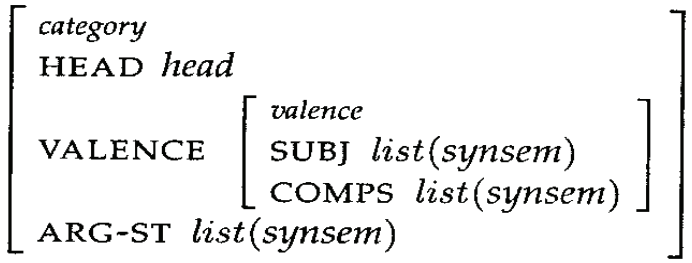

sign
PHON phon
SYNSEM
synsem
LOCAL
local
CAT category
CONT content
CONX context
NONLOCAL nonlocal
content
psoa
nom-obj
INDEX index
RESTR set(psoa)
love
LOVER index
LOVED index
give
GIVER index
GIVEN index
RECIPIENT index
chair
INST index
...

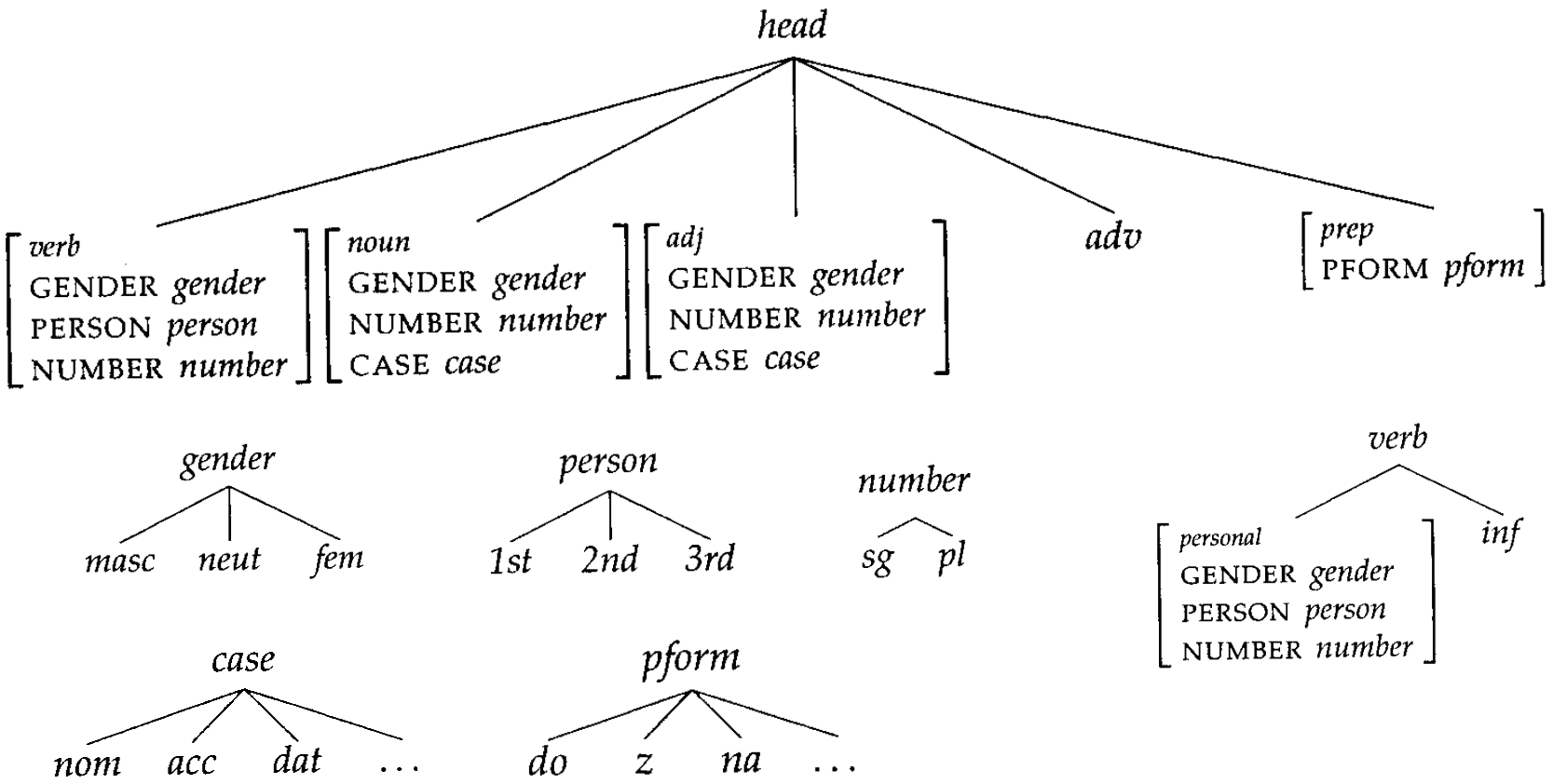
kochać:
$$\begin{bmatrix} \textit{love} \\ \text{LOVER } \boxed{1} \\ \text{LOVED } \boxed{2} \end{bmatrix}$$

dać:
$$\begin{bmatrix} \textit{give} \\ \text{GIVER } \boxed{1} \\ \text{GIVEN } \boxed{2} \\ \text{RECIPIENT } \boxed{3} \end{bmatrix}$$

krzesło:
$$\begin{bmatrix} \textit{nom-obj} \\ \text{INDEX } \boxed{1} \\ \text{RESTR } \left\{ \begin{bmatrix} \textit{chair} \\ \text{INST } \boxed{1} \end{bmatrix} \right\} \end{bmatrix}$$

znaczenie rzeczownika *krzesło* to „[1] takie, że zachodzi *chair*([1])", w skrócie „[1]: *chair*([1])"


category
HEAD head
VALENCE
valence
SUBJ list(synsem)
COMPS list(synsem)
ARG-ST list(synsem)

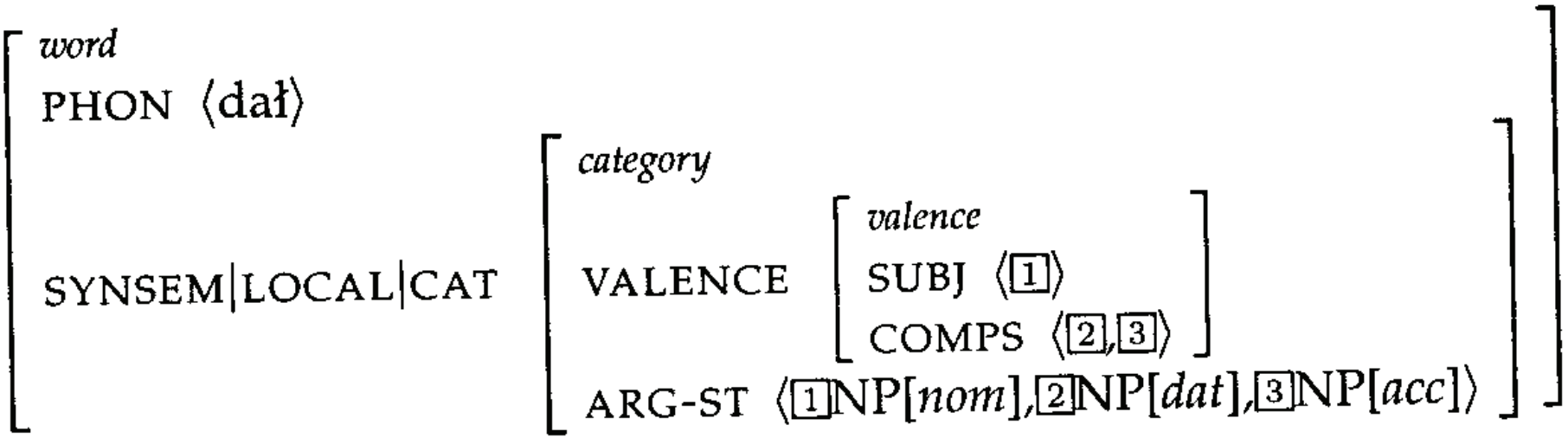

head
verb
GENDER gender
PERSON person
NUMBER number
noun
GENDER gender
NUMBER number
CASE case
adj
GENDER gender
NUMBER number
CASE case
adv
prep
PFORM pform
gender
masc neut fem
person
1st 2nd 3rd
number
sg pl
verb
personal
GENDER gender
PERSON person
NUMBER number
inf
case
nom acc dat ...
pform
do z na ...

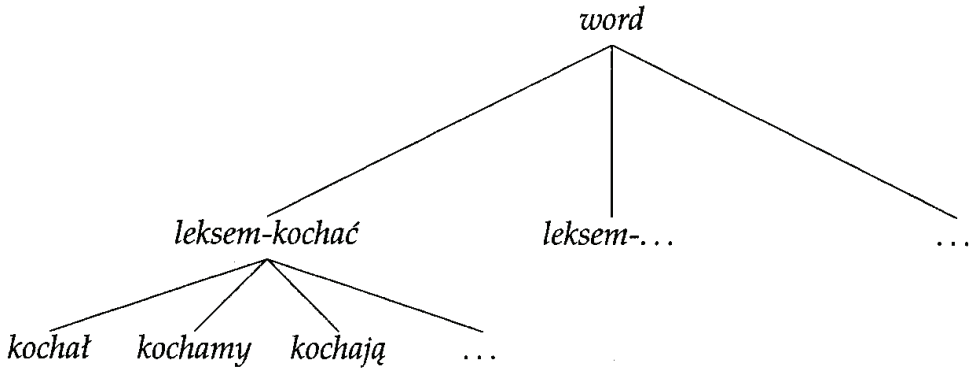
$$\left[\begin{array}{l} \textit{word} \\ \text{PHON } \langle \text{dał} \rangle \\ \\ \text{SYNSEM|LOCAL|CAT|VALENCE } \left[\begin{array}{l} \textit{valence} \\ \text{SUBJ } \langle \text{NP}[\textit{nom}] \rangle \\ \text{COMPS } \langle \text{NP}[\textit{dat}], \text{NP}[\textit{acc}] \rangle \end{array}\right] \end{array}\right]$$

$$\text{NP}[\textit{nom}] \stackrel{\text{df}}{=} \left[\begin{array}{l} \textit{synsem} \\ \\ \text{LOCAL|CAT } \left[\begin{array}{l} \textit{category} \\ \text{HEAD } \left[\begin{array}{l} \textit{noun} \\ \text{CASE } \textit{nom} \end{array}\right] \\ \text{VALENCE|COMPS } \langle\rangle \end{array}\right] \end{array}\right]$$

$$\begin{bmatrix} \textit{phrase} \\ \text{PHON} \ \langle \text{dał Marysi kwiaty} \rangle \\ \text{SYNSEM|LOCAL|CAT|VALENCE} \begin{bmatrix} \textit{valence} \\ \text{SUBJ} \ \langle \text{NP}[\textit{nom}] \rangle \\ \text{COMPS} \ \langle\rangle \end{bmatrix} \end{bmatrix}$$

$$\begin{bmatrix} \textit{phrase} \\ \text{PHON} \ \langle \text{Janek dał Marysi kwiaty} \rangle \\ \text{SYNSEM|LOCAL|CAT|VALENCE} \begin{bmatrix} \textit{valence} \\ \text{SUBJ} \ \langle\rangle \\ \text{COMPS} \ \langle\rangle \end{bmatrix} \end{bmatrix}$$


word
PHON ⟨dał⟩
SYNSEM|LOCAL|CAT
category
VALENCE
valence
SUBJ ⟨1⟩
COMPS ⟨2,3⟩
ARG-ST ⟨1NP[nom],2NP[dat],3NP[acc]⟩



a. *kochał* →

$$\begin{bmatrix} \text{PHON } \langle \text{kochał} \rangle \\ \text{SS|LOC|CAT|HEAD } \begin{bmatrix} \text{GENDER } \textit{masc} \\ \text{PERSON } \textit{3rd} \\ \text{NUMBER } \textit{sg} \end{bmatrix} \end{bmatrix}$$

b. *kochamy* →

$$\begin{bmatrix} \text{PHON } \langle \text{kochamy} \rangle \\ \text{SS|LOC|CAT|HEAD } \begin{bmatrix} \text{PERSON } \textit{1st} \\ \text{NUMBER } \textit{pl} \end{bmatrix} \end{bmatrix}$$

c. *leksem-kochać* →

$$\begin{bmatrix} \text{SS|LOC} \begin{bmatrix} \textit{local} \\ \text{CAT} \begin{bmatrix} \textit{category} \\ \text{HEAD } \textit{verb} \\ \text{VAL} \begin{bmatrix} \textit{valence} \\ \text{SUBJ } \langle \boxed{1}\text{NP}[\textit{nom}]_{\boxed{3}} \rangle \\ \text{COMPS } \langle \boxed{2}\text{NP}[\textit{acc}]_{\boxed{4}} \rangle \end{bmatrix} \\ \text{ARG-ST } \langle \boxed{1}, \boxed{2} \rangle \end{bmatrix} \\ \text{CONT} \begin{bmatrix} \textit{love} \\ \text{LOVER } \boxed{3} \\ \text{LOVED } \boxed{4} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

$$\text{NP}[\textit{nom}]_{\boxed{3}} \stackrel{\text{df}}{=} \begin{bmatrix} \textit{synsem} \\ \text{LOCAL} \begin{bmatrix} \textit{local} \\ \text{CAT} \begin{bmatrix} \textit{category} \\ \text{HEAD} \begin{bmatrix} \textit{noun} \\ \text{CASE } \textit{nom} \end{bmatrix} \\ \text{VALENCE|COMPS } \langle\rangle \end{bmatrix} \\ \text{CONT} \begin{bmatrix} \textit{nom-obj} \\ \text{INDEX } \boxed{3} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Zasada słownikowa

$$word \rightarrow \mathrm{HS}_1 \vee \mathrm{HS}_2 \vee \ldots \vee \mathrm{HS}_n$$

$$\mathrm{HS}_1 \stackrel{\mathrm{df}}{=} \left[\begin{array}{l} \textit{word} \\ \mathrm{PHON}\ \langle \text{kochał} \rangle \\ \mathrm{SS|LOC} \left[\begin{array}{ll} \textit{local} & \\ \mathrm{CAT} & \left[\begin{array}{ll} \textit{category} & \\ \mathrm{HEAD} & \left[\begin{array}{l} \textit{verb} \\ \mathrm{GENDER}\ \textit{masc} \\ \mathrm{PERSON}\ \textit{3rd} \\ \mathrm{NUMBER}\ \textit{sg} \end{array}\right] \\ \mathrm{VAL|SUBJ}\ \textit{nelist} & \\ \mathrm{ARG\text{-}ST}\ \langle \boxed{1}\mathrm{NP}[\textit{nom}]_{\boxed{3}}, \boxed{2}\mathrm{NP}[\textit{acc}]_{\boxed{4}} \rangle & \end{array}\right] \\ \mathrm{CONT} & \left[\begin{array}{l} \textit{love} \\ \mathrm{LOVER}\ \boxed{3} \\ \mathrm{LOVED}\ \boxed{4} \end{array}\right] \end{array}\right] \end{array}\right]$$

Zasada walencji

$$
\textit{phrase} \rightarrow \left[\begin{array}{l}
\text{SS|LOC|CAT|VAL} \left[\begin{array}{l}\textit{valence}\\ \text{SUBJ } \boxed{1'}\\ \text{COMPS } \boxed{2'}\end{array}\right]\\
\text{HD-DTR|SS|LOC|CAT|VAL} \left[\begin{array}{l}\textit{valence}\\ \text{SUBJ } \boxed{1}\\ \text{COMPS } \boxed{2}\end{array}\right]\\
\text{NONHD-DTRS } \boxed{3}
\end{array}\right]
$$

$$
\wedge\ ((\texttt{synsems-signs}(\boxed{2},\boxed{3}) \wedge \boxed{2'} = \textit{elist} \wedge \boxed{1} = \boxed{1'}) \vee (\texttt{synsems-signs}(\boxed{1},\boxed{3}) \wedge \boxed{1'} = \textit{elist} \wedge \boxed{2} = \boxed{2'}))
$$

`synsems-signs(<>,<>).`

`synsems-signs(<`$\boxed{1}$`|`$\boxed{2}$`>,<`$\boxed{1'}$`|`$\boxed{2'}$`>)` $\overset{\forall}{\Longleftarrow}$

$$
\boxed{1'} = \left[\begin{array}{l}\textit{sign}\\ \text{SYNSEM } \boxed{1}\end{array}\right]
$$

$\wedge$ `synsems-signs(`$\boxed{2}$`,`$\boxed{2'}$`).`

$$\textit{subj-phrase} \rightarrow \begin{bmatrix} \text{SS|LOC|CAT|VAL|SUBJ } \textit{elist} \\ \text{HD-DTR|SS|LOC|CAT|VAL} \begin{bmatrix} \textit{valence} \\ \text{SUBJ } \textit{nelist} \\ \text{COMPS } \textit{elist} \end{bmatrix} \end{bmatrix}$$

$$\textit{comps-phrase} \rightarrow \begin{bmatrix} \text{SS|LOC|CAT|VAL|COMPS } \textit{elist} \\ \text{HD-DTR|SS|LOC|CAT|VAL|COMPS } \textit{nelist} \end{bmatrix}$$

Zasada wartości PHON

$$\begin{bmatrix} \textit{phrase} \\ \text{PHON } \boxed{0} \\ \text{HD-DTR } \boxed{1} \\ \text{NONHD-DTRS } \boxed{2} \end{bmatrix} \rightarrow \texttt{signs-phons(<}\boxed{1}\texttt{|}\boxed{2}\texttt{>,}\boxed{3}\texttt{)} \wedge \texttt{permute(}\boxed{3}\texttt{,}\boxed{4}\texttt{)} \wedge \texttt{flatten(}\boxed{4}\texttt{,}\boxed{0}\texttt{)}$$

signs-phons(<>,<>).

signs-phons(<1|2>,<1′|2′>) $\overset{\forall}{\Longleftarrow}$

$\boxed{1} = \begin{bmatrix} \textit{sign} \\ \text{PHON } \boxed{1'} \end{bmatrix}$

$\wedge$ signs-phons(2,2′).

permute(<>,<>).

permute(1,<2|3>) $\overset{\forall}{\Longleftarrow}$

select(2,1,1′)

$\wedge$ permute(1′,3).

select(1,<1|2>,2).

select(1,<4|2>,<4|3>) $\overset{\forall}{\Longleftarrow}$

select(1,2,3).

flatten(<>,<>).

flatten(<1|2>,3) $\overset{\forall}{\Longleftarrow}$

flatten(2,2′)

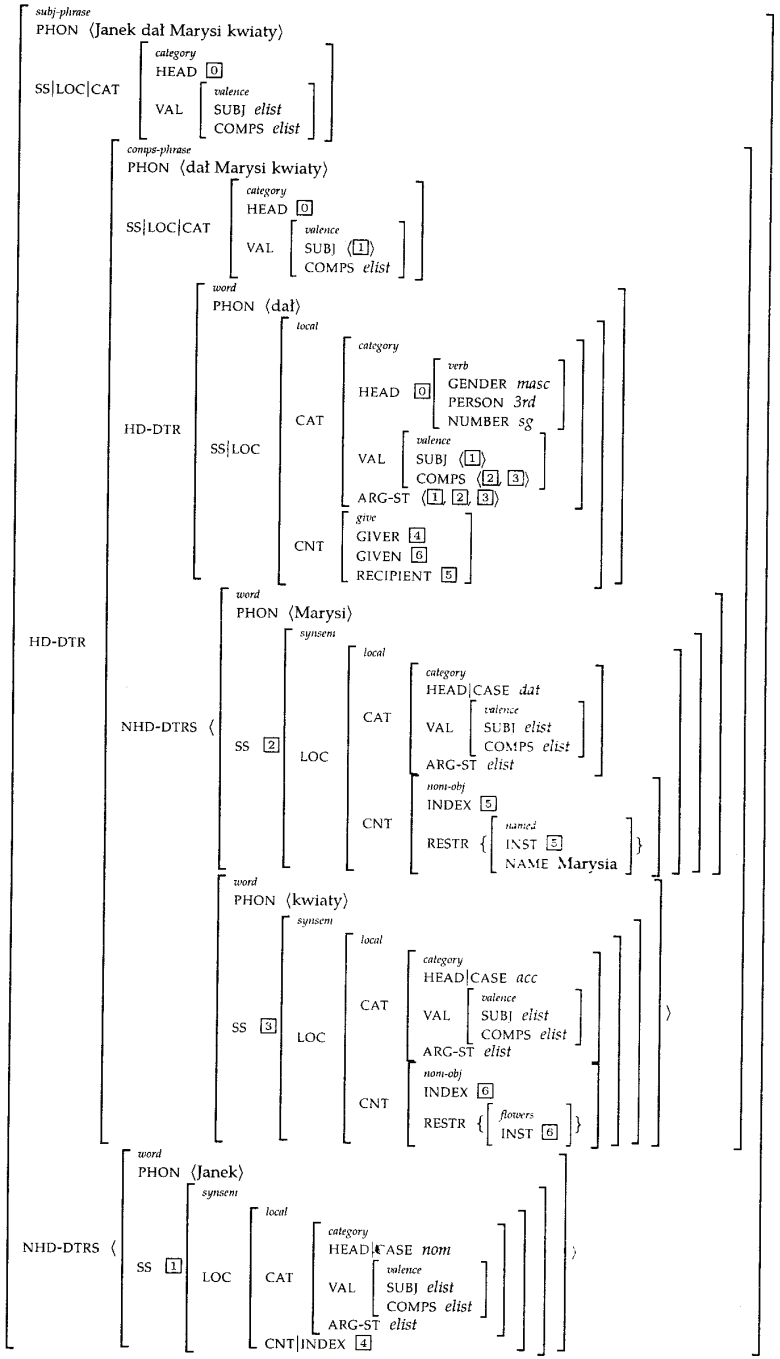
$\wedge$ append(1,2′,3).

$$\mathrm{HS}_d \stackrel{\mathrm{df}}{=} \left[\begin{array}{l} \textit{word} \\ \textsc{phon}\ \langle \text{dał} \rangle \\ \textsc{ss|loc}\ \left[\begin{array}{l} \textit{local} \\ \textsc{cat}\ \left[\begin{array}{l} \textit{category} \\ \textsc{head}\ \left[\begin{array}{l} \textit{verb} \\ \textsc{gender}\ \textit{masc} \\ \textsc{person}\ \textit{3rd} \\ \textsc{number}\ \textit{sg} \end{array}\right] \\ \textsc{val|subj}\ \textit{nelist} \\ \textsc{arg-st}\ \langle \boxed{1}\mathrm{NP}[\textit{nom}]_{\boxed{4}}, \boxed{2}\mathrm{NP}[\textit{dat}]_{\boxed{5}}, \boxed{3}\mathrm{NP}[\textit{acc}]_{\boxed{6}} \rangle \end{array}\right] \\ \textsc{cont}\ \left[\begin{array}{l} \textit{give} \\ \textsc{giver}\ \boxed{4} \\ \textsc{given}\ \boxed{6} \\ \textsc{recipient}\ \boxed{5} \end{array}\right] \end{array}\right] \end{array}\right]$$

$$\mathrm{HS}_j \stackrel{\mathrm{df}}{=} \left[\begin{array}{l} \textit{word} \\ \textsc{phon}\ \langle \text{Janek} \rangle \\ \textsc{ss|loc}\ \left[\begin{array}{l} \textit{local} \\ \textsc{cat}\ \left[\begin{array}{l} \textit{category} \\ \textsc{head}\ \left[\begin{array}{l} \textit{noun} \\ \textsc{case}\ \textit{nom} \end{array}\right] \\ \textsc{arg-st}\ \textit{elist} \end{array}\right] \\ \textsc{cont}\ \left[\begin{array}{l} \textit{nom-obj} \\ \textsc{index}\ \boxed{1} \\ \textsc{restr}\ \left\{\left[\begin{array}{l} \textit{named} \\ \textsc{inst}\ \boxed{1} \\ \textsc{name}\ \text{Janek} \end{array}\right]\right\} \end{array}\right] \end{array}\right] \end{array}\right]$$

$$\mathrm{HS}_m \stackrel{\mathrm{df}}{=} \left[\begin{array}{l} \textit{word} \\ \textsc{phon}\ \langle \text{Marysi} \rangle \\ \textsc{ss|loc}\ \left[\begin{array}{l} \textit{local} \\ \textsc{cat}\ \left[\begin{array}{l} \textit{category} \\ \textsc{head}\ \left[\begin{array}{l} \textit{noun} \\ \textsc{case}\ \textit{dat} \end{array}\right] \\ \textsc{arg-st}\ \textit{elist} \end{array}\right] \\ \textsc{cont}\ \left[\begin{array}{l} \textit{nom-obj} \\ \textsc{index}\ \boxed{1} \\ \textsc{restr}\ \left\{\left[\begin{array}{l} \textit{named} \\ \textsc{inst}\ \boxed{1} \\ \textsc{name}\ \text{Marysia} \end{array}\right]\right\} \end{array}\right] \end{array}\right] \end{array}\right]$$

$$\mathrm{HS}_k \stackrel{\mathrm{df}}{=} \left[\begin{array}{l} \textit{word} \\ \textsc{phon}\ \langle \text{kwiaty} \rangle \\ \textsc{ss|loc}\ \left[\begin{array}{l} \textit{local} \\ \textsc{cat}\ \left[\begin{array}{l} \textit{category} \\ \textsc{head}\ \left[\begin{array}{l} \textit{noun} \\ \textsc{case}\ \textit{acc} \end{array}\right] \\ \textsc{arg-st}\ \textit{elist} \end{array}\right] \\ \textsc{cont}\ \left[\begin{array}{l} \textit{nom-obj} \\ \textsc{index}\ \boxed{1} \\ \textsc{restr}\ \left\{\left[\begin{array}{l} \textit{flowers} \\ \textsc{inst}\ \boxed{1} \end{array}\right]\right\} \end{array}\right] \end{array}\right] \end{array}\right]$$


subj-phrase
PHON ⟨Janek dał Marysi kwiaty⟩
SS|LOC|CAT
category
HEAD 0
VAL
valence
SUBJ elist
COMPS elist
comps-phrase
PHON ⟨dał Marysi kwiaty⟩
SS|LOC|CAT
category
HEAD 0
VAL
valence
SUBJ ⟨1⟩
COMPS elist
HD-DTR
word
PHON ⟨dał⟩
SS|LOC
local
CAT
category
HEAD 0
verb
GENDER masc
PERSON 3rd
NUMBER sg
VAL
valence
SUBJ ⟨1⟩
COMPS ⟨2, 3⟩
ARG-ST ⟨1, 2, 3⟩
CNT
give
GIVER 4
GIVEN 6
RECIPIENT 5
HD-DTR
word
PHON ⟨Marysi⟩
NHD-DTRS ⟨
SS 2
synsem
LOC
local
CAT
category
HEAD|CASE dat
VAL
valence
SUBJ elist
COMPS elist
ARG-ST elist
CNT
nom-obj
INDEX 5
RESTR {
named
INST 5
NAME Marysia
}
word
PHON ⟨kwiaty⟩
SS 3
synsem
LOC
local
CAT
category
HEAD|CASE acc
VAL
valence
SUBJ elist
COMPS elist
ARG-ST elist
CNT
nom-obj
INDEX 6
RESTR {
flowers
INST 6
}
⟩
NHD-DTRS ⟨
word
PHON ⟨Janek⟩
SS 1
synsem
LOC
local
CAT
category
HEAD|CASE nom
VAL
valence
SUBJ elist
COMPS elist
ARG-ST elist
CNT|INDEX 4
⟩

Teoria deklaratywna

Odwracalne gramatyki komputerowe

Rozwiązywanie ograniczeń (*constraint solving*)