AUGMENTED X'-SCHEMES

Neculai CURTEANU

Institute for Computer Technique and Informatics [*]
IASI Branch, IASI
ROMANIA

Abstract

The paper presents the embedding of an original parsing strategy for Romanian, called Segmentation-Cohesion-Dependency (SCD) , into Chomsky's well-known Government and Binding (BG) theory. In order to bring closer the SCD concepts and techniques to the GB theory, the following questions have had to be dealt with: (1) a specification of the Principle of Maximal Projection (PMP), (2) extending the X'-theory (X - bar) by introdusing Augmented X' (AX')-schemes, these ones being obtained by (3) specific constraints imposed on the new shapes of X'-schemes. (4) The AX'-schemes can be represented in terms of a tree (paranthetic) language, whose tran-slation in a logic programming language follows naturally.

1. Introduction

Our approach represents a theoretical adaptation of Chomsky's subset of theories to the practical results obtained from and designed for SCD parsing strategy /Curteanu 1986/, /Curteanu 1987/. Mainly designed for Romanian, this strategy proved to be effective (with some peculiarities) also for English /Curteanu 1988b/. Let us consider the classical notations: X'-the second level of X projection, X''=XP-the third (top) level of projection for X=N,V,A,P. Consider also S , S' and PMP(X)=XP. We do not intend to discuss the problems of S''=PMP(S). We shall call X'-structures (middle level in the maximal projections) as "groups": a group contains an overt or an empty head, wrapped by its specifiers (denoted Specif) and modifiers (denoted Modif) but without argument(s) (denoted Arg). Thus X'=XG. NG=N' is called Elem-entary Noun Group (ENG). Tensed Verbal Group (TVG) is V' with the (morphologic or syntactic) feature of being tensed. Non-tensed Verbal Group (NVG) is V' with the (morphologic) feature of being untensed. ECP abbreviates the Empty Category Principle. We note that X'-structures are naturally obtained by SCD rules.

As Chomsky remarks "...the distinction between modifiers and arguments is notoriously difficult in certain cases" /Chomsky 1982 p.44/. Giving a general solution for the dis-tribution of the verbal nature modifiers , the AX'-schemes offer also the correct assign-ment of the functional dependencies between the head and its Modif, on the one hand, and the head and its Arg(s), on the other hand. The paper presents a PMP specification,which the adopted solutions are based on. It mainly says that not only V has subcategorization properties, but also some N and A (part of these verbal Ns or As being, in some lang-uages, V). The new features of the AX'-schemes are: (1) an arbitrary number $(n \geqslant o)$ of Arg(s) are permitted, (2) the AP is always a Modif and (3) the PP is always an Arg in AX'-schemes

[*] Mailing Address: Universitatea "Al.I.Cuza",
Seminarul Matematic "A.Myller",66oo IASI,
ROMANIA

whose overt or empty head is N,V,A. The sub-categorization properties depend on an ab initio assigment, at the dictionary level, of the lexical feature VERB of the category. (4) The overt or empty (PRO) subject is considered as a special argument of the tensed (S) or un-tensed (NP) maximal projections of V. (5) In the hypotheses (4) and (2) the traditional , formal VP is dissolved into S or complex, ver-bal NP. (6) The phenomena of binding and bounding are easier remarked and solved this way.

2. A Specification of PMP

We have shown in /Curteanu 1987/, /Curt-eanu 1988a/ that, within the SCD parsing strategy, there are detached basic (middle level) structures of the form X' (like ENG, TVG, NVG). They are elements of the (tensed or untensed) clause relation predicate - arguments. For all these X'-structures we could find the same functional representation shape, written as a LISP tree. The new prop-osed AX'-schemes have their origin in the governing idea of the SCD strategy: the (max-imal) use of the verbal character of, tradit-ionally non-verbal (including for Romanian) categories, N and A. Let us consider the fol-lowing PMP specification:

> The subcategorization properties of the (consistent) syntactic categories N,V,A depend on the following lexical and, res-pectively, morphological features of these categories: VERBality, with the values ACT and EXIST, and PREDicativity, with the values T and NIL.

Such a specification of the maximal project-ion function is necessary because in many languages (including Romanian) the verbal quality of the non-verbal categories like N or (sometimes) A has to be discovered and assigned at the dictionary level. E.g., for English, the nouns that "verbalize" in -ing have, of course, the VERB feature. There are, also, quite few participles which are not adjectives and fewer adjectives which are not participles, despite their VERB quality (and thus, of the subcategorization property).

3. Augmented X' - Schemes

Before proposing the new AX'-schemes, we simply remark a (more adequate, at least for our approach) alternative shape to the class-ical X'-scheme. This should be:
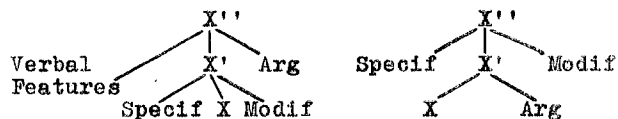
Modified shape instead of the Classical one(C)

Fig.3.1.   Fig.3.2.

In what follows, for convenience, we still use the classical shape (C).

The general form of the AX'-scheme is:

Left column

```
              X''
         ┌─────┼─────┐
      Specif  X'    Modif
          ┌───┼────────┐
         X  VERB Arg₁ Arg₂..Argₙ
```

X=N, V, A, P
Augmented X'-scheme (A1)
Fig.3.3.

```
              [X'']ᵢ
         ┌─────┼─────┐
      Specif [X']ᵢ  Modif=APᵢ
         ┌────┼────────┐
        [X]ᵢ VERB Arg₁ Arg₂..Argₙ
```

X=N, V, A, e
AX'-scheme for AP (A2.1)
Fig.3.4.

```
              [X'']ᵢ
         ┌─────┼─────┐
      Specif [X']ᵢ  Modif
         ┌────┼──────────┐
        [X]ᵢ VERB Arg₁..Argᵢ=PP..Argₙ
```

X=N, V, A, e
AX'-scheme for PP (A2.2)
Fig.3.5.

```
              S'
         ┌────┼────┐
      COMP       S
              ┌───┼────────┐
           VG PRED Arg₁ Arg₂..Argₙ
```
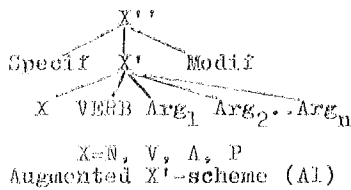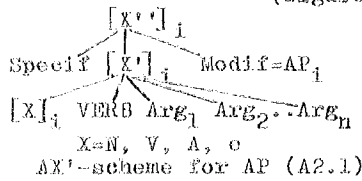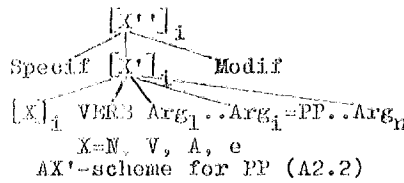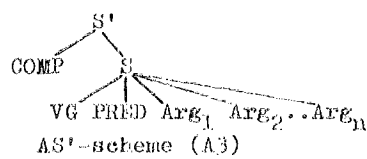
AS'-scheme (A3)
Fig.3.6.

the classical S'-scheme. If VG:=TVG (PRED=T)
S becomes the classical (finite) clause. If
VG:=NVG, (PRED=NIL), S becomes an infinite
clause, naturally assimilated by schemes
(A2.1) and (A2.2) to a complex, verbal NP. In
this way, the value X=V in (A1) is purely
theoretical since the maximal projection of
V is embedded into S or NP.

The theoretical arguments for the AX' -
schemes are:
a1. The special position of the (grammatical)
subject in a S'-scheme is, in fact, an A- or
a Case position. It becomes a right θ - pos-
ition only after syntactic and semantic
calculi.
a2. At the syntactic level, the subject pos-
ition presents well-known ambiguities.
a3. In /Wehrli 1984/, as well as in lexical-
functional grammars /Sells 1985/ and unific-
ation-based grammars /Shieber 1986/, there
are adopted similar points of view. This
option does not represent any impediment to
apply specific mechanisms, e.g. control theo-
ry, to obtain the A- or θ - position of the
overt or empty subject.
a4. Such a solution has as consequence the
extension of the maximal projection of VG to
(tensed or untensed) sentencial shapes, ent-
ailing the direct transition to the logical
form and analysis of natural language, a
profit incomparably greater than the formal
losing (embedding into S or NP) of the VP.
a5. The binding and bounding phenomena are
better revealed and translated. E.g., we are
dealing, as bounding nodes, with tensed and
untensed S.

Here there are some **practical arguments.**

Right column (Remark text, set beside figures)

Remark.The formal
differences bet-
ween (C) and (A1)
are: the occur -
ence of the lexic
-al feature VERB
and the arbitrary
number of Arg(s).
Let us further
specify (A1)
(figures 3.4. and 3.5.)

The feature
value ACT of
the categories
N and A (and
implicitly V)
is assigned
when these
categories
correspond to
complex, act-
ional verbs.
The value
EXIST is
implicitly or
explicitly in-
troduced by
the forms and
meanings of
the verb
"to be".
For S we con-
sider the
AS'-scheme
(Fig.3.6.).
The feature
PRED is
similar to the
abstract categ
-ory INFL from

Section 4

4. Augmented X'-Schemes at Work

There are underlined the words having
assigned the feature value ACT,and used the
SCD segmentation rules.

Ex.4.1./floare albă /de cireş/
Ex.4.1./the cherry /white blossom/

```
                    NP
           ┌─────────┼──────────────┐
        Specif  N'                Modif=AP
           │    ┌────┼─────┐    ┌────┼──────┐
         the  N EXIST Arg=NP  Specif A'   Modif
              │           │    NIL  A EXIST Arg  NIL
          blossom  Specif N' Modif     white   NIL
                      │    ┌────┼───┐
                    N EXIST Arg NIL
                      │          │
                   cherry       NIL
```
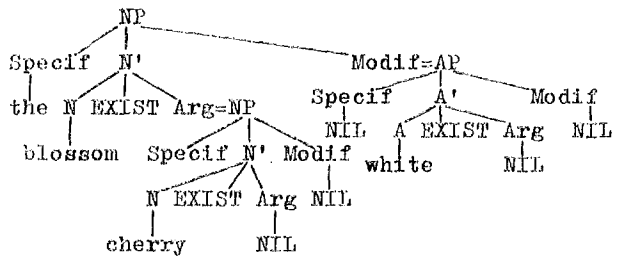
Fig.4.1.

Ex.4.2./floare albă/,/bătută/ /de vînt/
Ex.4.2./white blossom/,/blown/ /by the wind/

```
                 NPᵢ
         ┌────────┼──────────┐
      Specif     N'ᵢ       Modif=AP
         │    ┌────┼────┐       │
       the  Nᵢ EXIST Arg     white
         blossom      NIL
```

Fig.4.2.1.

```
                 [NPᵉ]ᵢ
       ┌──────────┼──────────────────────┐
   Specif[Nᵉ]ᵢ                         Modif=AP
   NIL [Nᵉ]ᵢ EXIST Arg     Specif  A'        Modif
              NIL       NIL  NIL A ACT Arg=PP   NIL
                                  │     ┌──┼──┐
                                blown   P  NP
                                        by the wind
```
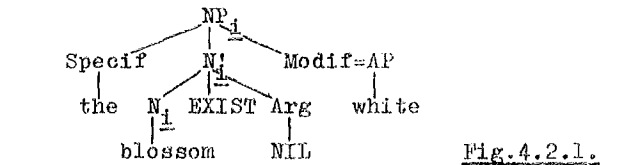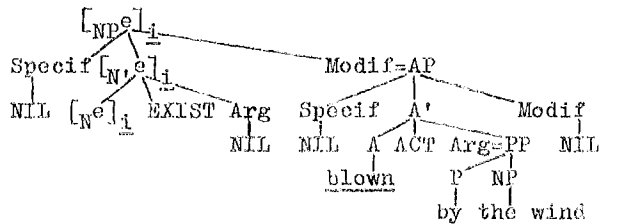
Fig.4.2.2.

Ex.4.3./educat/[de tatăl său] /corespunzător/
/cu vechile principii/
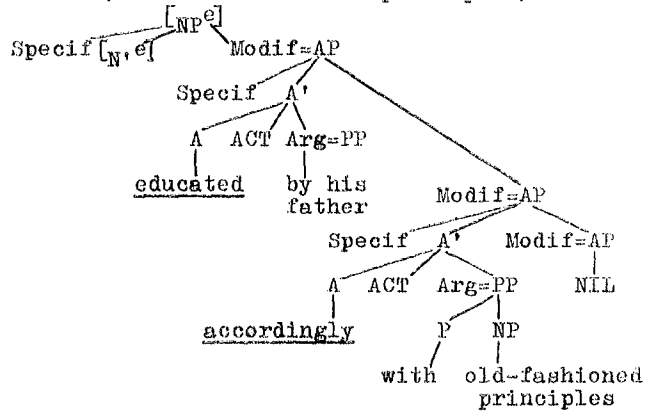Ex.4.3./educated /[by his father] /accordingly/
/with old-fashioned principles/

```
                 [NPᵉ]
         ┌──────────┼──────────────┐
   Specif[Nᵉ]               Modif=AP
                         ┌─────┼──────────┐
                      Specif  A'
                         ┌─────┼──────┐
                        A   ACT  Arg=PP
                        │         │
                     educated  by his
                               father      Modif=AP
                                       ┌──────┼──────┐
                                    Specif   A'     Modif=AP
                                       ┌─────┼───┐    NIL
                                      A   ACT Arg=PP
                                      │         │
                                  accordingly  P  NP
                                               with old-fashioned
                                                    principles
```

Fig.4.3.

Ex.4.4./Omulᵢ//căruiaᵢ/PROⱼ/iᵢ-am dat /
/cartea//PROⱼ a părăsit/ ţara.//
Ex.4.4./The manᵢ//whomⱼ/ Iⱼ/gave/the book/
/PROᵢ left/ the country.//

Specif N'
the N_i EXIST Arg Modif=S'
man_i NIL COMP
whom_i
S
VG PRED Arg_1 Arg_2 Arg_3
gave T T_j e_i the book

Fig.4.4.1.

S
VG PRED Arg_1 Arg_2
left T PRO_i the country   Fig.4.4.2.
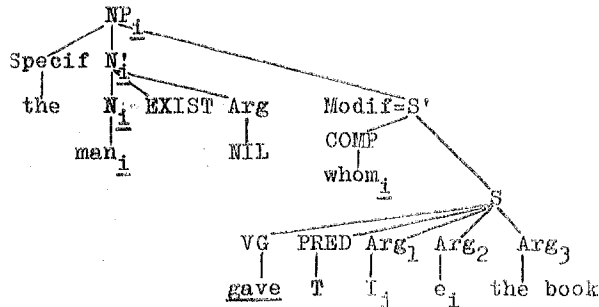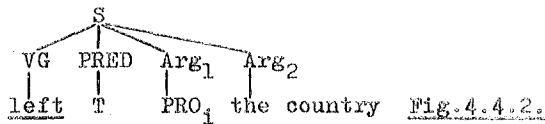
Ex.4.5./Ion_i/ a convins-o_j /pe Maria_j/ /că

deplasarea_k /ei_j/ la Bucureşti/este

utilă e_k / Universităţii.//

Ex.4.5'./John_i/convinced/ Mary_j //that /her_j

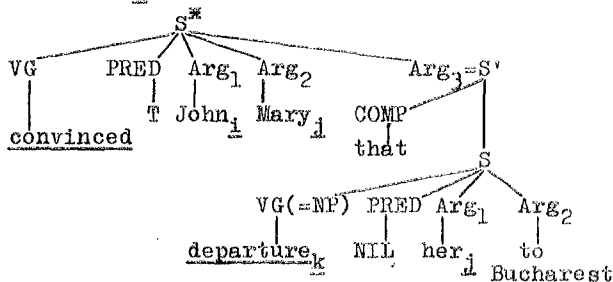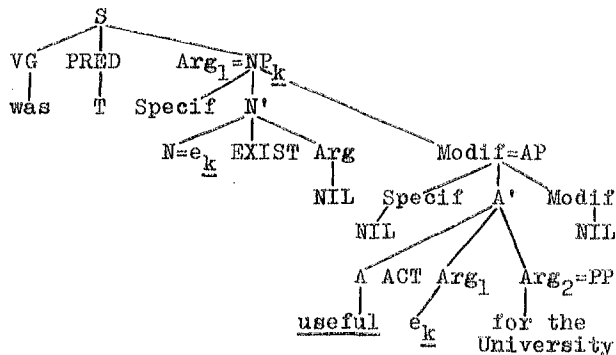/departure_k/to Bucharest//was useful

e_k /for the University.//

S*
VG PRED Arg_1 Arg_2 Arg_3=S'
convinced T John_i Mary_j COMP
that
S
VG(=NP) PRED Arg_1 Arg_2
departure_k NIL her_j to
Bucharest

Fig.4.5.1.

S
VG PRED Arg_1=NP_k
was T Specif N'
N=e_k EXIST Arg Modif=AP
NIL Specif A' Modif
NIL NIL
A ACT Arg_1 Arg_2=PP
useful e_k for the
University

Fig.4.5.2.

By a raising transformation we obtain the alternative "small clause":

S
VG PRED Arg_1=NP_k Arg_2=PP
was-useful T

Fig.4.5.3.

Remark. In /Chomsky 1982/ there are discussed the various aspects of such constructions. We simply observe that the choice between the last two structures depends on the implementation structures of the logic representation language.
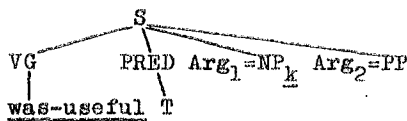
5. Final Remarks

The main constrains on the AX'-schemes are:
c1. The (abstract) feature VERB actions at the lexical level. It takes the values ACT or EXIST according to the verbal-actional or (implicit or explicit) verbal-existential nature of its head.
c2. The (abstract) feature PRED is subsumed by the classical INFL. It takes the values T and NIL. Other features, specific to INFL (like AGR), could be added.
c3. The assignment of the Modif role to AP (in particular to NVG) is justified both by the ECP and the identic behaviour of AP in the frame of ENG and TVG-NVG.
c4. The assignment of the Arg role to PP is justified both by the ECP particularization to NP and by the PMP specification for the consistent categories N,V,A. P is considered to be "non-consistent" since, as a head, it cannot be properly "modified" (and "specified") but only "argumentized".
c5. The VP (in fact, VG) embedding into S or NP was already justified in Section 3.
c6. The AX'-scheme (A2.1) offers a very suitable frame for the ellipsis solving.

The specific inter- and intrapropositional structures (ENG, NVG) and relations are properly revealed by the SCD strategy. The verbal-actional or verbal-existential nature of the heads of such (middle level) X'-structures has to be functionally reflected. The AX'-schemes could represent good premises for the maximality level of the verbal functionality use in syntax, hence in the lexical semantics. This has been the purpose of our paper.

References

Chomsky, Noam (1982) 'Lectures on Government and Binding' Dordrecht, Foris Publications.
Curteanu,N. & O.Popârda & I.Oprea (1986) 'Algorithme d'analyse du groupe nominal (Domain roumain)' Revue Roumaine de Linguistique, XXXI, No.2, p.171-19o, Bucarest.
Curteanu,N.& O.Popârda & I.Oprea (1987) 'An Algorithmic Description of the Behaviour of Noun Groups' Computers and Artificial Intelligence,Vol.6,Nr.1,p.17-34, Bratislava.
Curteanu,N.& O.Popârda (1988) 'X'-schémas de representation de modificateurs de nature verbale' Revue Roumaine de Linguistique, XXXIII, No.1, p. 45-54 , Bucarest.
Curteanu,N. et al. (1988) 'ISATER-Interactive System for Assisted Translation: English - Romanian' Research Contract, I.C.T.I. Iaşi (in preparation).
Hajičova,E. & P.Sgall (1986) 'The Ordering Principle' Prague Bulletin of Mathematical Linguistics, Nr.45, p.1-26.
Marcus,S.(Ed.) (1983) 'Contextual Ambiguities in Natural and Artificial Languages' Ghent, Belgium.
Saint-Dizier,P. (1986) 'An Approach to Natural Language Semantics in Logic Programming' J. of Logic Prog., Vol.3, Nr.4, p.329-356.
Sells,P. (1985) 'Lectures on Contemporary Syntactic Theories' CSLI Lect. Notes Nr.3, Stanford, California.
Shieber,S. (1986) 'Introduction to Unification-Based Approaches to Grammar' CSLI Lect. Notes Nr.4, Stanford, California.
Wehrli,E. (1984) 'A Government-Binding Parser for French' Univ.de Geneve,Working Paper 48.