Approaches to Hungarian **1** (1985): 79–92

The internal structure of Noun Phrases

András Kornai
Institute of Linguistics, Hungarian Academy of Sciences

0 Introduction

The aim of this paper is to describe the *internal* syntax of Hungarian noun phrases in terms of their immediate constituent structure. The external syntax of NPs, i.e. their distribution in larger phrases and sentences will be mentioned only in passing – for a more detailed discussion, see Kenesei (1985). Section 1 deals with the lower bar-levels, and Section 2 describes the possessive constructions: the results are summarized as context-free rule schemata in Section 3.

1 The easy parts of the NP

A noun preceded by an adjective is, perhaps, the most simple combination of words that can reasonably be called an NP in Hungarian. (Nouns followed by modifiers are discussed in Laczkó (1985)). Yet a simple phrase-structure rule, such as

(1) NP → A N

is clearly insufficient for the description of this trivial fact. For instance, the case of the nouns that can be substituted for such NPs must be the same as the case of the head noun

Piroska	még	nem	evett	paprikás	gulyást
Piroschka	yet	not	eat-PAST	spicy	goulash-ACC
'Piroschka has not yet eaten spicy goulash'					
*Piroska	még	nem	evett	paprikás	gulyás
Piroschka	yet	not	eat-PAST	spicy	goulash-NOM

and the above rule makes no provision for this. In order to capture the regularities of constituent structure in a phrase-structure formalism, it is thus necessary to employ rules containing complex symbols of some sort. The following discussion is intended primarily to make the use of such symbols clear. Most rules will be only sketched first, and their rigorous statement is deferred to Section 3.

Since other constituents (numerals, quantifiers, etc.) can be added to A+N constructions, they are not maximal projections, so the use of "NP" on the left-hand side of (1) is somewhat misleading. The usual solution to this problem is the introduction of bar-levels (cf. e.g. Harris (1951, ch 16), Jackendoff (1977, ch 3)). In the Jackendovian scheme, we must take it into account that the adjective is

optional, and assign the construction to the first bar level. This would give us a rule like[1]

(1A) $N^1 \rightarrow (A)\ N$

But if we follow the Harrisian principles of level-assignment, then the first thing we note is that the substitution of A+N for N can be repeated as many times as we wish. Indeed, constructions with stacked adjectives like

kicsi	puha	sárga	francia	párna
'little	soft	yellow	french	pillow'

are quite common, and this means we should adopt the rule

(1B) $N \rightarrow (A)\ N$

rather than rule (1A). Since Hungarian NPs can contain several obligatory constituents, it would be impossible to use only rules permitted by the Uniform 3-Level Hypothesis of Jackendoff (1977), and in what follows, non-repeatable substitutions (and only these) will correspond to raised bar-level. However, I will use a rule like

(1C) $N^1 \rightarrow (AP)\ N$

to describe the position of the adjectival noun-modifiers: readers bothered by this inconsistency should subtract one from every bar level (other than zero) throughout the paper.

The only major departure from the Harrisian mode of immediate constituent analysis is the adoption of words, rather than morphemes, as basic syntactic units. As a consequence, the category symbols will have to be annotated for the paradigmatic form of the word in question. This will be achieved by a liberal use of morphosyntactic features, such as αCASE (where α ranges over NOM, ... , FOR), the person-number features ±ME, ±YOU, ±PL, etc. These features are all inflectional (in the sense of Anderson 1982) – for their morphological realization see Kálmán (1985). With the aid of complex symbols made up from the basic category symbol (such as N, V, Num, etc.), a number indicating bar-level, and the morphosyntactic features (which will be always enclosed in angled brackets), (1C) can be replaced by the rule schema (1D). (For ANP, see Section 2.)

(1D) N< 1 αCASE βPL γPOS δANP>→
A <n -CASE -PL -POS -ANP> N< 0 αCASE βPL γPOS δANP>

Thus, every morphosyntactic feature behaves here as a head feature (in the sense of GPSG, see. e.g. Gazdar 1982); and the adjective modifying the head noun does *not* agree with it in number, case, or any other feature. For the adjunction of numerals, a similar rule can be stated:

(2) N< 2 αCASE -PL βPOS γANP>→
Num<n -CASE -PL -POS -ANP> N< 1 αCASE -PL βPOS γANP>

[1] The final version of the rules is lettered F, G, ...

The scope of this paper does not permit a detailed discussion of the numeral phrase or of the adjectival phrase. Nevertheless, it should be mentioned here that Num< 1 > can contain a measure term:

három	kiló	rothadt	alma
three	kilogram	rotten	apple

and that numerals (or NumPs) and adjectives (or APs) cannot be interchanged:

*rothadt három kiló alma
*A< k > Num< k > N

Rule (2) makes it clear that NPs containing a numeral must be singular:

*három almák
*Num N<+PL>

so in order to provide for an unbroken projection line for plural NPs, we might add the rules

(2A) N< 2 αCASE βPL γPOS δANP>→ N< 1 αCASE βPL γPOS δANP>

(2B) N< 1 αCASE βPL γPOS δANP>→ N< 0 αCASE βPL γPOS δANP>

The definite article *a/az* attaches proclitically to the N< 2 >: if there is a numeral (and an adjective), it has to precede the numeral,

a	három	rothadt	alma
the	three	rotten	apple
*három	a	rothadt	alma
*Num	Art	N< 1 >	

and if there is an adjective, but no numeral, the article has to precede the adjective:

*rothadt az alma
*A Art N (qua NP)

Since only one article can appear in any N< 2 >, all the above facts can be conveniently summarized in the rule

(3) N< 3 >→ Art N< 2 >

The verb in Hungarian has two forms for every permissible tense/mood/aspect/person/number combination (see Kálmán (1985)). In intransitive constructions, the so-called 'subjective' forms are selected, and the forms traditionally called 'objective' appear only in transitive constructions with definite objects.[2] In other words, the verb agrees with the object (more precisely, with

[2] This picture is somewhat complicated by the suffix *-lak/-lek* which appears only in the first person singular paradigm. Second person objects always select this suffix, no matter whether they are definite (e.g. *mindnyájatok* 'you all') or indefinite (e.g. *bennetek* 'you (pl)') with -ME or +PL subjects.

the NP having accusative marking) *with respect to definiteness:* this makes it particularly easy to test the definiteness of NPs in Hungarian. The definite article, of course, will make the resulting construction definite: this will be captured by adding the feature +D to the left-hand side of (3). This feature is inherent on proper nouns:

Látom	Attilát	*Látok Attilát
see+D	Attila-ACC	*V<-D> N<+D ACC>
'I see Attila'		

In colloquial Hungarian, the definite article can also be added,

Látom	az	Attilát
see+D	the	Attila-ACC

so there is an additional rule

(3A) N< 3 +D>→ Art N<+D>

The cooccurrence of numerals and adjectives with proper nouns is largely restricted to metaphorical usage:

a magyar Beethoven
the Hungarian B.

Certain proper nouns cannot appear with definite article:

*a Lisszabon *az Afrika

(There is a tendency in normative grammars to put every personal name in this class.) Adjectival constructions show this to be a purely local phenomenon:

a gyönyörű Lisszabon	az éhező Afrika
'the beautiful Lisbon'	'the hunger-stricken Africa[3]'

2 Possessive constructions

There are three kinds of possessive constructions in Hungarian, two of them syntactic, and one morphological. The latter is purely anaphoric: the presence of the suffix *-é* on a head noun refers to something in the possession of the head:

Ödöné
Ed's ...

This *anaphora possessiva* (ANP) suffix is *not* a genitival ending. Firstly, the ordinary genitive construction is absent from Hungarian:

*Ödöné könyv	*Ödöné a könyv
*N<+ANP> N	*N<+ANP> Art N (qua NP)

and secondly, case endings arein complementary distribution, while *-é* can cooccur

[3] There is also a "The Hague" class of proper names, the items of which must appear with the definite article.

with every case.

*Ödönneket	Ödönét
*N-DAT-ACC	'Ed-ANP-ACC'

With the aid of slash categories (in the original sense of GPSG, cf. Gazdar 1982), possessive anaphoric constructions can be related to NPs 'with an N missing'. In order to familiarize the reader with this kind of complex symbol, let us first investigate constructions like

a három sárgát	nyolccal	azt
the three yellow-ACC	eight-INS	the-ACC

Such expressions are best translated to English with anaphoric *one:* 'the three yellow ones', 'eight ones', 'the one'. These expression occur, by and large, in the same sort of coordination reduction contexts as in English:

Én	a	fekete	kocsival	megyek,	ő	pedig	a	pirossal
I	the	black	car-INS	go-PRES	he	but	the	red-ACC

'I'll go in the black car and he in the red one'

(For more detailed discussion, see Kerkovits (1985)) There are, then, NPs with the head missing: the rules

(4A) N< 1 > /N → (AP) N/N
(4B) N< 2 > /N → (NumP) N/N
(4C) N< 3 > /N → Art N/N

are parallel to the rules in (1), (2), and (3). N/N is a phonetically null noun with no phonological effect whatsoever: its case, number, etc. features are spelled out on the preceding element. With the aid of this notation,

(5) N<+ANP>=N/N<POS>

The expression on the left-hand side of the equation has the same semantic content as the expression on the right-hand side. +POS is the same possessive suffix that appears in syntactic possessive constructions like

Ödön	könyve	Ödönnek	a	könyve
Ed	book-POS	Ed-DAT	the	book-POS

Let us first take the simpler of these two, where the possessor is in the nominative form. Since the possessed noun can be modified by numerals and/or adjectives, but cannot take the definite article,

Ödön	két	érdekes	könyve	*Ödön	a	könyve
Ed	two	interesting	book-POS	*N	Art	N<POS>

'Ed's two interesting books'

the head of the construction is an N< 2 >. The substitution cannot be iterated:

*János	Péter	könyve(je)
*N	N	N<POS> (<POS>)

so the whole construction is on a higher bar-level. This gives us the rule

(6) N< 3 >→ N< 3 -POS> N< 2 +POS>

The three bars of the first NP on the right-hand side (i.e. the possessor) can be motivated as follows. First, the article can appear in the leftmost position of these:

a	fiu	kutyája
the	boy	dog-POS

'the boy's dog'

and second, this article forms a constituent with the following N, as can be seen from the non-existence of such constructions as

*a	Lisszabon	főutcája
*Art	Lisbon	main-street-POS

(cf Section 1. above), and, further, from a comparison of the obviative article *a/az* with the proximate article *e*. In constructions like

e	fiu	kutyája
this	boy	dog-POS

e has narrow scope, and if we suppose that the article belongs to the whole construction, the nearness of the boy and the possible remoteness of the dog remains unexplained.

The possessive construction is definite even if the head noun is indefinite, and it is -POS in spite of the fact that the head noun appears in +POS form. The former observation means that possessive NPs can only be substituted by definite nouns in object position, and the latter means that they can be substituted by nouns that do not have possessive suffixation:

Látom	Péter	kalapját	*Látom	kalapot
see-1st-sg-D	Peter	hat-POS-ACC	*N<+D>	N<-D>

'I see Peter's hat'

Piroska	gulyása	paprikás	* (a) gulyása	paprikás
Piroschka	goulash-POS	spicy (is)	* (Art) N<+POS>	N

'Piroschka's goulash is spicy'

Case, ANP, and PL behave like head features here. This is also true of the other possessive construction, which involves dative marking on the possessor, and must have an N< 3 > in the position of the possessed element:

*Ödönnek	könyve
*N<+DAT>	N<+POS>

Since the substitution is, again, nonrepeatable,

*a fiunak a lánynak a könyve
*N< 3 +DAT> N< 3 +DAT> N< 3 +POS>

the dative possessive construction must be assigned a higher bar-level:

(7) N< 4 >→ N< 3 +DAT> N< 3 +POS>

The possessor is at least N< 3 >, as can be seen from the constructions

e	fiunak	a	könyve	*a	Lisszabonnak	a	főutcája
the	boy-DAT	the	book-POS	*Art	N<-ART>	Art	N<+POS>
(near)	(near)	(far)					

where -ART is used simply as a diacritic distinguishing those nouns that do not take the article as sister constituent. The status of iterated possessive constructions such as

% a	fiunak	a	barátjának	a	könyve
the	boy-DAT	the	friend-POS-DAT	the	book-POS

'the book of the friend of the boy'

% a	fiu	barátja	könyve
the	boy	friend-POS	book-POS

'the book of the friend of the boy'

is not quite clear: for those speakers, who accept them, we might take the possessor to be N< 4 >. The general acceptance of iterated possessives starting with an interrogative pronoun

kinek	a	barátjának	a	könyve
who-DAT	the	friend-POS-DAT	the	book-POS

'whose friend's book'

seems to support this conclusion. Either way, the present assignment of bar-levels makes it possible to generate iterated constructions like

a	fiu	barátjának	a	könyve
the	boy	friend-POS-DAT	the	book-POS

'the boy's friend's book'

where the dative possessor is a nominative possessive construction, while the reverse construction, where the nominative possessor is a dative possessive construction, is not permitted:

*a	fiunak	a	barátja	könyve
*Art	N<DAT>	Art	N<POS>	N<POS>

This is in accordance with the (in these cases, unequivocal) judgments of native speakers.

It should be noted here that quantifiers, with the sole exception of *minden* 'every', have roughly the same distribution as numerals. —it Minden, however, is in complementary distribution with the definite article

*a minden könyv *minden a könyv (qua NP)

and, indeed, can be taken as an indefinite article[4]

[4] Traditional grammars usually call the numeral *egy* 'one' the indefinite article, although its distribution is almost identical to that of other numerals. Its peculiarities stem from the fact that it has positional variants after articles: **az egy* is replaced by *az egyik* 'the one' or by *az egyetlen* 'the only' and **minden egy* is replaced by *minden egyes* 'every single'.

Látok	minden	könyvet	*Látom	minden	könyvet
see-1st-Sg-D	every	book-ACC	*V<+D>	every	N<ACC>

'I see every book'

If the possessor is a personal pronoun, only the nominative construction can be used:

az	én	könyvem	*nekem	a	könyvem[5]
the	I	book-POS	*I-DAT	Art	N<+POS>

'my book'

The situation here is parallel to that in subject-predicate constructions. The possessive affix agrees with the possessor in person and number, and the pronoun can be dropped:

a	te	könyved	a	könyved
the	Pro2ndSg	book-POS-2nd-Sg	the	book-POS-2nd-Sg
'your	book'		'your	book'

The article, however, has to be retained in most of the cases:

*ő	könyve	*könyve
*Pro	N<+POS>	*N<+POS>

This fact is captured by subsuming these pronouns under Art< 1 +D> as optional complements:

(8A) N< 3 +D -POS>→ Art< 1 +D> N< 2 +POS>
(8B) Art< 1 +D>→ Art< +D> (Pro)

Finally, the parallel between N<+ANP> and N N/N<+POS> can be extended to pronouns. The possessive suffixes *enyém, tiéd, ...* 'mine, yours, ...' can be treated like Pr<+ANP>: this is clear from possessive sentences like

a	könyv	Ödöné	volt	a	könyv	az	enyém	volt
the	book	Ed-POS	was	the	book	the	mine	was
'the	book	was	Ed's'	'the	book	was	mine'	

Now, if we substitute Pro N/N<POS> for Pro<+ANP> we get

a	könyv	az	én	N/N-m
the	book	the	I	N/N<+POS>
'the	book	is	my	(book)'

[5] The dative can be used if the possessor and the possessed item are not in the same construction:

Nekem	a	könyveim	sikeresek
I-DAT	the	book-PL-POS	successful-PL

'My BOOKS are successful'

and in fact it must be used if the possessor is not adjacent to the possessed item:

Nekem	sikeresek	a	könyveim
I-DAT	successful-PL	the	book-PL-POS

'My books are successful'

3 The rules

(1F) N< 1 αPOS βPL γANP δCAS ϵD>→
(A< n -POS -PL -ANP -CAS>) N< 0 αPOS βPL γANP δCAS ϵD>

(1G) N< 1 αPOS βPL γANP δCAS ϵD>/N →
(A< n -POS -PL -ANP -CAS>) N< 0 αPOS βPL γANP δCAS ϵD>/N

(2F) N< 2 αPOS -PL βANP γCAS δD>→
(Num< n -POS -PL -ANP -CAS>) N< 1 αPOS -PL βANP γCAS δD>

(2G) N< 2 αPOS +PL βANP γCAS δD>→ N< 1 αPOS +PL βANP γCAS δD>

(2H) N< 2 αPOS -PL βANP γCAS δD>/N →
(Num< n -POS -PL -ANP -CAS>) N< 1 αPOS -PL βANP γCAS δD>/N

(2I) N< 2 αPOS +PL βANP γCAS δD>/N →
N< 1 αPOS +PL βANP γCAS δD>/N

(3F) N< 3 αPOS βPL γANP δCAS ϵD>→
Art< ϵD> N< 2 αPOS βPL γANP δCAS ±D>

(3G) N< 3 αPOS βPL γANP δCAS +D>→
N< 0 αPOS βPL γANP δCAS +D>

(3H) N< 3 αPOS βPL γANP δCAS ϵD>/N →
Art< ϵD> N< 2 αPOS βPL γANP δCAS ±D>/N

(5F) N< 3 αPOS βPL ANP< γ PL> δCAS ϵD>=
N< 3 αPOS βPL -ANP -CAS ±D> +
+ N< 2 +POS γPL -ANP δCAS ϵD>/N

(6F) N<3 -POS αPL βANP γCAS +D>→
N<3 ±POS ±PL -ANP -CAS ±D> N<2 +POS αPL βANP γCAS ±D>

(7F) N<4 -POS αPL βANP γCAS +D>→
N<3%4 ±POS ±PL ±ANP +DAT ±D> N<3 +POS αPL βANP γCAS +D>

(8F) N<3 -POS αPL βANP γCAS +D>→Art<1 +D δME ϵYOU ζPL>
N<2 POS< δME ϵYOU ζPL> αPL βANP γCAS ±D>

(8G) N<3 -POS αPL βANP γCAS +D>→Art<0 ±D>
N<2 POS< ±ME ±YOU ±PL> αPL βANP γCAS ±D>

(8H) Art<1 +D αME βYOU γPL>→Art<+D> Pro< αME βYOU γPL>

References

Anderson, S. R. 1982: Where is morphology? Linguistic Inquiry 13, 571-612
Gazdar, G. 1982: Phrase structure grammar. In: Jacobson-Pullum (eds): The Nature of Syntactic Representation. Reidel, Dordrecht
Harris, Z. 1951: Methods in Structural Linguistics. University of Chicago Press
Jackendoff, R. 1977: X-bar Syntax. MIT Press, Cambridge, Mass.
Kálmán, L. 1985: Inflectional Morphology, in this volume
Kenesei, I. 1985: Subordinate Clauses, in this volume
Kerkovits, A. 1985: Ellipsis, in this volume
Laczkó, T. 1985: Deverbal nouns and their complements, in this volume