Asymmetries in Hungarian

LÁSZLÓ MARÁCZ
(Gröningen)

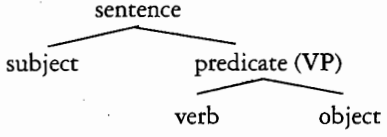
1. THE CONFIGURATIONALITY ISSUE

In this chapter, I will introduce the *configurationality* issue in general (cf. section 1.1.), and discuss this question with respect to Hungarian in particular (cf. section 1.2.).[1]

1.1. On Configurationality in General

Before going into details, let me first sketch in short the *core problem* of the configurationality issue. In many languages there is direct evidence for a special grammatical relation of *subject* and a different one of *object.* Syntactic rules may apply to a combination of the object and verb but not to a combination of the subject and verb. Compare, for example, the VP-rules, such as VP-preposing and VP-deletion, in English. Languages with such rules are said to display subject-object *asymmetries.* The appearance of subject-object asymmetries in a particular language is considered as evidence for the different positions of the subject and the object in the structure of the sentence in that language. If we abstract away from surface word order variation, this may be represented in the following tree diagram:

(1)



In some languages, there seems to be little or no evidence available for this subject-predicate partitioning of the sentence. Subject-object asymmetries seem to be missing in these languages. The question arises whether these languages still differentiate subjects and objects in a fundamental way.

While generative grammarians had taken it for granted that in English there is a syntactic VP-node, and had devised a series of constituency tests to show that there is a subject-predicate partitioning of the sentence, some linguists discovered that the

(1) See also the introduction of Marácz and Muysken (1989) for a historical overview of the configurationality debate, discussion of some proposals, and methodological questions concerning configurationality.

[*ASJU*, XXIV-2, 1990, 407-524]

Ed. note: The publication of this work will be completed in the next two issues of *ASJU*. Bibliographical references will be published in the last issue.

tests did not carry over easily to non-Indo-European languages. Arguments for a VP-node were hard to come by in those languages. An example of this is Hinds (1974) who argued that there was no reason to assume a VP-constituent for Japanese.[2]

Similarly, syntacticians found it difficult to reconcile the considerable freedom of word order in some languages with the mechanism of phrase-structure rules. It was assumed that phrase-structure rules generated ordered strings of elements only. An example is Staal's (1967) work on Sanskrit.[3] Staal argued that the order of subject, verb and object was completely free in Sanskrit, and he proposed to replace the formalism of ordered trees of Chomsky (1965; 1977) by that of 'wild' or unordered trees. These trees indicated to what constituent a given element belongs but not the order of elements within that constituent. Note that Staal did maintain a VP-node in Sanskrit. Staal's proposal was, however, not a theoretical improvement, because as Chomsky (1965: 123-127) argued, set-systems are equivalent to concatenation-systems.[4]

The main impulse for work on configurationality came in the late seventies, when Ken Hale discovered that aboriginal Australian languages such as Warlpiri were hard to classify in terms of typological notions current until then. He observed that Warlpiri allows an extremely free word order, that is, any ordering of constituents will yield a grammatical sentence. The only restriction on word order in that language is that the auxiliary verb (Aux) must be in second position:

(2) Kurdungku ka maliki wajilipinyi
child-ERG Aux-pres dog-ABS chase-nonpast
Maliki ka kurdungku wajilipinyi
Maliki ka wajilipinyi kurdungku
Wajilipinyi ka kurdungku maliki
Wajilipinyi ka maliki kurdungku
Kurdungku ka wajilipinyi maliki
'The child is chasing the dog.'
(Hale 1981: 1)

Hale (1981) observed further that the extreme freedom of word order is not only restricted to the verbal arguments but may also involve constituents which are a single semantic unit corresponding to NP in English. Note that the parts of the phrase *two small children* in Warlpiri is an instance of a 'split' constituent. Compare:

(3) *Kurdujarrarluk* kapala maliki wajilipinyi *witajarrarlu*
child-dual-ERG Aux-pres-dual dog-ABS chase-nonpast small-dual-ERG
Maliki kapala *kurdujarrarlu* wajilipinyi *witajarrarlu*
Witajarrarlu kapala maliki wajilipinyi *kurdujarrarlu*
(etc., any order with Aux in second position)
'The two small children are chasing the dog.'
(Hale 1981: 1)

(2) In those days even some reseachers of Germanic languages did not assume a VP. See, for example, the treatment of V-raising in Dutch by Evers (1975).

(3) See also Šaumjan and Soboleva's (1963) study on free word order in Russian. They argued that the phrase marker of Russian could be captured more easily by an unordered set-system instead of a concatenation-system.

(4) Chomsky (1965: 123-127) acknowledges, however, that freedom of word order cannot be captured in terms of the theory of transformations at that time.

Combining the insights of Šaumjan and Soboleva (1963), Staal (1967), and Hinds (1974), Hale (1981; originally written in 1978) proposed to capture these observations by defining the basic syntactic structures of Warlpiri by the following minimal rule:

(4) $E \rightarrow W^*$

This rule states that in Warlpiri expressions (E) are formed by stringing words (W) together. Hale, unlike for example Staal in Sanskrit, did not assume the presence of a syntactic VP-node in the phrase-structure of Warlpiri.

In Hale (1980), the typological distinction between free and fixed word order languages conformed to the formalism of X'-theory as outlined in Chomsky (1970) and developed in Jackendoff (1977). The X'-scheme generates the following endocentric rules:

(5) a. $X'' \rightarrow \ldots X' \ldots$
b. $X' \rightarrow \ldots X \ldots$ (where X is N, V, ...)

According to Hale (1981), some languages employ both (5a) and (5b) for the realization of their endocentric categories, the *configurational* languages, whereas the syntax of *non-configurational* languages contains only rule (5b). Rule (5b) expresses three things: (i) Each endocentric category has a head, (ii) the order of modifiers is free, and (iii) constituents are 'flat' in that there is no intermediate structure between a head and its maximal projection.[5]

Hale (1982) suggested that the difference between configurational and non-configurational languages is not only restricted to fixed versus free word order. Rather, there is a clustering of so-called non-configurational properties. Hale listed the following 'diagnostics':

(6) a. 'Free' word order
b. The use of split or discontinuous constituents
c. Free or frequent *pro*-drop
d. The lack of NP-movement
e. Lack of expletive elements (like *it, there,* etc.)
f. Use of a rich case-system
g. Complex verb words
h. The lack of VP-rules (like VP-preposing, VP-deletion, etc.)
i. The lack of ECP-effects[6]

Hale argued that some of these properties (such as the lack of standard ECP-effects and *pro*-drop) could be derived by assuming that in non-configurational languages, i.e. languages with one-prime categories, the notion government, defined as a relation between a head and its direct sister, is absent. It turned out, however, that this list of diagnostics could not characterize the type. Languages classified as non-

(5) Rule (5b) may also specify the relative order of heads and complements. For example, the fact that heads in Japanese are category-final can be expressed as follows (Japanese was analysed at that time as a non-configurational language, see Hale 1980 and Farmer 1980):
(i) $X' \rightarrow \ldots X$

(6) Diagnostic (6i) has been added by Huang (1982).

configurational displayed at most only a subset of these properties. For example, Hungarian and Japanese, which were characterized in the literature as non-configurational, do not possess a 'strong' Aux-node such as Warlpiri, or Navajo (cf. 6g). Furthermore, established configurational languages such as Italian or Dutch may also display a subset of the non-configurational characteristics. For example, Italian has 'free' word order, free or frequent *pro*-drop, and lack of ECP-effects with long Wh-movement (cf. Rizzi 1982). Dutch exhibits 'free' word order, *pro*-drop with non-referential expressions, lack of VP-rules, and lack of ECP-effects with long Wh-movement (cf. Koster 1986). Hence, it became less clear what the 'proper' diagnostics of a non-configurational language were.

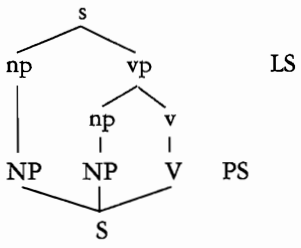
In the course of this study, I will demonstrate that Hungarian displays a subset of the diagnostics of non-configurationality, and that these phenomena may be accounted for *without* assuming a non-configurational phrase-structure for its syntax. It will be argued that they may be attributed to independently motivated principles of UG and properties of Hungarian. 'Free' word order is not so free after all. More and more phenomena have been found which restrict freedom of word order. Hungarian has even neutral word orders (cf. chapter two). Overt expletives are lacking but there is some evidence that non-overt expletives may be present (cf. chapter four). Some VP-rules such as VP-preposing may apply under specific circumstances in Hungarian as well (cf. chapter five). The occurrence of split constituents is heavily restricted both syntactically and semantically in Hungarian (cf. chapter four). Free or frequent *pro*-drop falls under the *Pro*-drop Parameter (cf. chapter four). The lack of NP-movement follows from the way θ-roles are related to syntactic structure in Hungarian (cf. chapter three). The lack of standard VP-rules such as VP-deletion is due to the syntactic properties of I in Hungarian (cf. chapter five). Finally, the lack of ECP-effects with long Wh-movement is dependent on the fact that the minimal maximal domain of the subject in Hungarian happens to coincide with that of the object, namely the CP (cf. chapter five).

The modular approach to grammar narrowed the distinction between the grammars of configurational and non-configurational languages. It initiated the search for *subject-object asymmetries* in non-configurational languages. A reasonable hypothesis, then, was that evidence for subject-object asymmetries would turn up in the modules of the grammar. Hale (1983) discovered subject-object asymmetries within the domain of binding theory (reflexive-reciprocae binding) and control theory in Warlpiri. Notice that after the discovery of subject-object asymmetries in non-configurational languages the term 'non-configurational' was no longer a particularly appropriate one. Therefore, the configurationality puzzle shifted from the problem of free word order to the following question. How is the cluster of *both* subject-object symmetries (see, for example (6h), the lack of VP-rules) and subject-object asymmetries in the grammar of a particular language to be accounted for? An initial answer to this question was suggested in Chomsky (1981).

Chomsky assumed that all languages are configurational at Lexical Structure (LS), a subpart of D-structure, which is an abstract, mobile structure representing the hierarchical organization of a predicator and its direct arguments, but not at the overt categorial representation, called Phrase Structure (PS). This latter representation was

assigned a flat structure in non-configurational languages. As a consequence, the phrase marker of a sentence in a non-configurational language was represented at each level of representation as a dual non-isomorphic syntactic structure. Consider:

(7)



Chomsky related the dichotomy between LS and PS in non-configurational languages and the isomorphicity between LS and PS in configurational languages to a parametrization of the Projection Principle. He hypothesized that in non-configurational languages the Projection Principle holds only at LS, i.e. 'Assume a Grammatical Function (GF)', whereas in configurational languages it holds of the pair (LS, PS). This approach accounted for some of the properties of non-configurational languages.

Free word order was handled by free lexical-insertion and base-generation at PS, subject-object symmetries were attributed to PS, the representation where the subject and object are equally prominent, and subject-object asymmetries were attributed to LS, where a hierarchical division of the arguments of the verb is made. Chomsky's parameter *Assume GF* has been elaborated in more detail by Hale (1983), Mohanan (1983) and Zubizarreta and Vergnaud (1982).[7]

Note that a relaxation of the Projection Principle led to an *anomaly* in the theory of UG. Firstly, Chomsky (1981) redefined the core of the generative research program. The theory of phrase-structure grammars was eliminated from the theory of UG, and was replaced by new core principles such as the Projection Principle. By parametrizing the Projection Principle, Chomsky created an internal conflict in this research program. Secondly, note that a representation like (7) is not a reduced phrase marker in the sense of Lasnik and Kupin (1977), because not every pairs of nodes dominates or precedes the other in a single phrase marker. Chomsky (1982: 14) states: "It should be clear that the theory of phrase-structure has no standing as a component of UG". From this, we may conclude that there is no theoretical objection against representations like (7). This would, however, imply that the theory of reduced phrase markers should be given up. Certainly, an undesirable step.

In reaction to such rather radical proposals, other researchers working in the generative tradition have proposed to account for cases of apparent free word order with mechanisms that remain much closer to the standard assumptions of generative grammar. In these approaches to the configurationality puzzle, researchers tried to account for this typological split by parametrizing a subcomponent of the grammar.

(7) This idea of double representation led also to the extensive study of the formal properties of phrase markers. See Zubizarreta and Vergnaud (1982), Higgingbotham (1985), and Speas (1986).

Stowell (1981) suggests that relaxing the adjacency condition on Case-assignment has the effect of allowing for free word order. In fixed word order languages, the object, for example, has to remain next to the verb because in those languages Case-assignment requires adjacency. If the object were anywhere else, it would not be Case-marked, leading to an ungrammatical result. Consequently, if there is no adjacency requirement on Case-assignment in a language, the order of elements can be much freer.

Van Riemsdijk (1982) interprets Hale's observations on Warlpiri in terms of the difference between the syntactic representations most familiar to us and phonological representations. Warlpiri clauses would have no tree structure but they would be organized phonologically. They would be subject to adjacency conditions of phonology rather than those of syntax.

In Saito (1982) and much related work, the assumption is made that in a free word order language such as Japanese the phrase-structure rules create a VP-node and ordered constituents, but that the possibility of freely adjoining constituents to the clause they are part of has the effect of allowing free word order.

Jelinek (1983) and Speas (1986) provide empirical evidence against a parametrization of the Projection Principle. They argue that even in Warlpiri and Navajo, the Projection Principle is satisfied by fully referential clitic pronouns that serve as verbal arguments. Therefore, they conclude that the Projection Principle is satisfied at all levels of representation even in non-configurational languages. Note that such a theory is in fact a notational variant of a theory which assumes a VP-node and the application of adjunction rules. The linking of 'dislocated' NPs in non-A-positions to the clitic pronouns in the A-positions of Aux is equivalent with the binding of A-positions by NPs which are in non-A-positions by the application of adjunction.[8]

Kuroda (1987) has proposed the Forced Agreement Parameter in order to derive the main typological differences between English and Japanese:

(8)

	English	Japanese
Visible Wh-movement	+	-
Scrambling	-	+
Topic-prominence	-	+

The Forced Agreement Parameter states that complements and heads in English, unlike in Japanese, must display agreement. As a consequence, the subject NP in [Spec, IP] in English, contrary to Japanese, must agree with the head of IP, i.e. I(nfl). The presence of an NP in this position blocks movement from the [Spec, VP] to the [Spec, IP] in English. Hence, the lack of scrambling in that language. In Japanese, on the other hand, nothing prevents the movement of an NP from [Spec, VP] to [Spec, IP]. This yields, then, scrambling in Japanese.

In sum, it seems to me that the configurationality puzzle consists of the study of the internal structure of the clusters of subject-object asymmetries and symmetries,

(8) Such a state of affairs happens more often as Chomsky (1981: 346) notes: "It is quite possible that alternative approaches that appear superficially to be quite different may fall together, when the proper level of abstraction is identified and clarified."

their relation and their position in a theory of UG. The focus of research has shifted from the superficial diagnostics of (6) towards the position of these clusters in a theory of UG. There are two possibilities to approach these questions:

Scenario I

The asymmetries are taken as the unmarked cluster, that is, they are generated by the subcomponents and principles of UG. The presence of this cluster in a particular grammar is taken as an indication that all languages are configurational, and consequently have a VP-node. This represents the *null-hypothesis,* because of the existence of established configurational languages such as English. The puzzle for proponents of this position can be phrased as follows: How is the cluster of symmetries covered in a theory of UG?

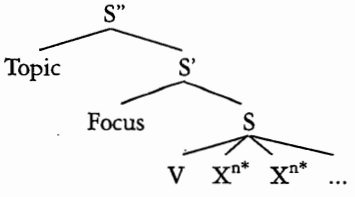
Scenario II

The cluster of symmetries is the unmarked one in the sense that it is generated by the phrase-structure of a particular language. This option is problematic from the point of UG. In the light of the existence of uncontroversially configurational languages it is rather *ad-hoc.* The questions to answer for proponents of this position are the following: What is the position of the cluster of symmetries in a theory of UG, and how is the cluster of asymmetries to be accounted for in the grammar of a particular language?

1.2. Configurationality and the Grammar of Hungarian

Much work in Hungarian syntax deals with the position of Hungarian with respect to the Configurationality Parameter. As we will point out in the next chapter, Hungarian allows 'free' word order. This, taken together with the absence of the most direct evidence for a configurational phrase-structure, has led some researchers to classify Hungarian as a non-configurational language. This position has been most clearly defended in the studies of É. Kiss (cf. É. Kiss 1981a, and subsequent literature). According to É. Kiss, the propositional part of the sentence is flat. She distinguishes between non-A-positions hierarchically ordered on the 'periphery' of the sentence (Topic, and Focus) and A-positions in S, and claims that move-α affects arguments by shifting them to any of the two peripheral positions in (1):

(1)



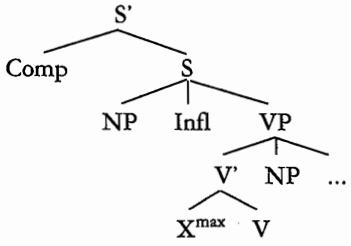
Topic may contain several maximal major categories, while Focus is restricted to a single constituent. Furthermore, the phrases in S may be scrambled. This structure lacks a VP-constituent. Hence, the subject and object have the same distribution

structurally. This hypothesis lead to the discovery of subject-object symmetries in Hungarian where in English asymmetries appear (cf. É. Kiss 1982a). These subject-object symmetries occur in several modules, like X'-theory (position of sentence adverbs, and absence of VP-rules), Wh-module (the lack of superiority effects), and quantification module (Topicalization of universal quantifiers). I will return in chapter five to an extensive discussion of subject-object symmetries and their position in Hungarian syntax. Note that the approach just discussed has a somewhat crude empiricist flavor.[9]

This approach supposes that hypotheses about how to cover variation in word order and the presence of subject-object symmetries should spring directly out of the way the data initially look to the investigator. É. Kiss proposes to account for the properties of Hungarian by postulating a special type of phrase structure, namely, the one depicted in (1). Since languages like English lack variation in word order and subject-object symmetries, they are assumed to have a quite distinct grammar (cf. É. Kiss 1982a; 1987c).

An alternative configurational approach to Hungarian syntax in line with scenario I above has been proposed in Horvath (1981; 1986a).[10] Horvath regards Hungarian as having a basic SVO-order and exhibiting much of the configurational character of, say, English in the operation of NP-movement, and hierarchical clause structure. A D-structure like (2) must be affected by various instances of move-α, including movement to Comp, Topicalization, Subject Postposing (a VP-adjunction rule), downgrading movement, scrambling, movement in LF, and so on in order to produce all the possible varieties of strings of constituents:

(2)



The X^{max} under V' provides for various preverbal constituents in neutral sentences such as verbal prefixes, determinerless nouns, predicative adjectives, and so on, and is vacated if some other constituent of the sentences is to occupy that position. This constituent will then receive Focus-interpretation at LF. Although the assumption of a VP-node represents the null-hypothesis and is thus preferred over the more impressionistic approach, Horvath's elaboration faces several problems. Elsewhere, I

(9) Gazdar et al. (1983: 5) refer to this type of approach as 'neo-empiricism'.

(10) In the literature, there are other pairs of competitive analysis concerning the configurational status of one and the same language. For example, a configurational versus non-configurational analysis has been proposed by: Eguzkitza (1986), Ortiz de Urbina (1986), and Salaburu (1985) versus Abaitua (1985), Azkarate et al. (1981), and Rebuschi (1985) for Basque; Den Besten (1982), Fanselow (1987), Koster (1986), and Webelhuth (1985) versus Halder (1985; 1986), Sternefeld (1984), and Tappe (1982) for German; Saito and Hoji (1983) versus Farmer (1980), Farmer et al. (1986), and Hale (1980) for Japanese; and Choe (1985) versus Yang (1982; 1984) for Korean.

have attempted to demonstrate that some of her actual tests on subject-object asymmetries suffer from empirical and theoretical shortcomings (cf. Marácz 1988a).

The following tests are incomplete, including the distribution of sentence adverbs, Quantifier Float, VP-deletion, and the distribution of bound pronouns (Weak Crossover effects). Other tests make the wrong predictions under the theoretical assumptions adopted by Horvath, like Subject Postposing as an instance of VP-adjunction, mixed configurational categories, Quantifier Float, and Weak Crossover effects (WCO). I will return to a more extensive discussion of these VP-tests and their theoretical consequences in the sections 4.6., 5.3., and 5.4.

An initial justification for the approach dictated by the null-hypothesis came from the observations made in Horvath (1981: 210) and É. Kiss (1982). They noted that subject-object asymmetries in Hungarian occur in the domain of WCO and reflexive binding. Since then the list of subject-object asymmetries has rapidly grown, involving various other modules of the grammar. I will catalogue these subject-object asymmetries in chapter five. The problem of Hungarian syntax became not only how to account for variation in word order but also how to account for a cluster of subject-object asymmetries and symmetries in the grammar?

Following Chomsky's (1981) suggestions on configurationality, some of these questions were tackled in Marácz (1986a). A serious disadvantage of the approach to assign the phrase marker of a sentence a dual representation (cf. 1.1.(7)), is, as I pointed out above, that it involves a drawback from a theoretical point of view. It leads to an relaxation of the theory of UG.[11] On the other hand, this hypothesis stimulated the following lines of research. Firstly, it initiated the search for subject-object asymmetries in Hungarian. Secondly, it made it necessary to reconsider the question of the mapping between LS and PS. These lines of research led to the discovery of other subject-object asymmetries in the domain of binding theory unambiguously showing that subject and object have a different distribution in Hungarian as well.

Elsewhere (cf. Marácz 1987c), I have proposed that the mapping of LS onto PS in Hungarian has the following four properties (cf. Koster 1987):

(3) a. Obligatoriness b. Biuniqueness
c. Identity d. Locality

Obligatoriness is supported by the fact that all lexically selected verbal arguments are present at surface structure. Hence, no lexically selected arguments may be lost during the derivation. The relation between LS and PS is biunique, that is, each argument at LS corresponds to exactly one constituent at PS. The relation between LS and PS is subject to an identity requirement involving either structural positions or morphological markers. Furthermore, the relation between LS and PS obeys a general locality constraint, i.e., the PS-constituent (or its place holder) appears in the domain of the verb whose LS contains the argument to which it is related. An approach which assumes that the mapping between LS and PS is restricted by the

(11) Compare also Horvath (1987) for this point.

properties in (3) is a notational variant of the theory which falls under scenario *I* of the preceding section. In that case, the VP-node remains *visible* at all levels of representation. In chapter four, I will discuss the mapping between LS and syntax in more detail and the status of the Projection Principle in Hungarian.

Although I think that enormous progress has been made in recent years, a number of empirical and theoretical questions concerning the syntax of Hungarian remain controversial. A more careful examination than hitherto is, in my view, required to account for freedom of word order, the properties of phrase-structure, and the parallel presence of clusters of subject-object asymmetries and symmetries. In the chapters to follow, I wish to make a modest contribution with respect to the settlement of these questions. The grouping of phenomena and their analysis is motivated by the theory of UG outlined in the introduction. It is my hope that this study will contribute to the definite solutions of the puzzles discussed above, and so will yield a deeper insight into the structure of Hungarian and the structure of UG.

2. THE PHRASE STRUCTURE OF HUNGARIAN

In the following sections, I will develop a theory for the *phrase structure* of Hungarian. Recently, some generalizations on word order in Hungarian have been observed. In section 2.1., I will discuss these generalizations. Furthermore, in section 2.2., these generalizations will serve as the basis for a theory of Hungarian phrase structure.

2.1. Descriptive Generalizations on Word Order in Hungarian

Hungarian has traditionally been regarded as a free word order language.[1] This freedom of word order is, however, only restricted to the sentential clause. As I will demonstrate below, other maximal projections, like NP and PP, have a fixed word order. In this section, I will consider some phenomena that are related to the question of word order. These phenomena have in common that they put restrictions on the sentential *word order* variation involving (*A*) the neutral word order, (*B*) linear restrictions on complex verb constructions, (*C*) the fixed Focus-position, (*D*) the position and interpretation of quantifiers, and (*E*) linear and hierarchical restrictions on maximal projections other than the clause. Let us consider (*A*) first.

(A) In the literature on Hungarian word order, there is no general agreement on the question whether Hungarian has a *neutral* sentence-order. The position that Hungarian lacks a neutral word order has been defended in É. Kiss (1981a; and subsequent literature), whereas Kiefer (1967), Horvath (1981; 1986), and Kálmán et al. (1986) hold the opposite view. It seems to me that the position in the latter references is the correct one. Hence, I will assume that Hungarian has a neutral sentence-order, namely, SVO. Let us consider, however, the position of É. Kiss first.

É. Kiss claims that the sentence has no neutral order, and that the only fixed part is constituted by the verb (cf. É. Kiss 1987: 39). The postverbal constituents may be scrambled around freely. In preverbal position, É. Kiss distinguishes two consecutive

(1) The following studies on word order in Hungarian have been undertaken by, among others, Ackerman and Komlósy (1983), Deák (1988), Dezső (1965), Horvath (1986a), Hunyadi (1985), Kálmán (1985a; 1985b), Kálmán et al. (1986), Kenesei (1985c; 1986b), Kiefer (1967; 1970), É. Kiss (1986b; 1987a), and Pléh (1982).

types of categories associated with different structural positions: An unstressed constituent, anda constituent immediately preverbal bearing primary stress. She presents the following taxonomy of word order variation in a transitive sentence (' indicates primary stress, *János* 'John', *Máriát* 'Mary-ACC', *szereti* 'loves'):

(1)	*I*	*II*	*III*	*IV*
	János	'Máriát	szereti	
	János Máriát		'szereti	
	Máriát	'János	szereti	
	Máriát János		'szereti	
	János		'szereti	Máriát
		'János	szereti	Máriát
	Máriát		'szereti	János
		'Máriát	szereti	János
			'Szereti	János Máriát
			'Szereti	Máriát János

(É. Kiss 1987: 39)

É. Kiss (1987: 39) claims further that this grouping of the complements has also a semantic-communicative function. The constituents in position *I* present the information which is known both to the speaker and hearer. According to her, these constituents possess Topic (T) function. Therefore, she calls this position the Topic-position. The constituent in position *II* bears primary stress, and it displays a fall in pitch. É. Kiss claims that this constituent is not only phonologically but also semantically the most prominent constituent of the sentence. It is the so-called Focus (F) of the sentence. Therefore, she calls this the Focus-position.[2]

Contrary to É. Kiss, Kiefer (1967), Horvath (1981; 1986), and Kálmán et al. (1986) claim that sentences with a neutral order do exist in Hungarian. The order in such sentences is *SVO*. Kálmán et al. (1986: 130), for example, distinguish the following two types of sentences in Hungarian: (*i*) *Corrective* sentences, and (*ii*) *neutral* sentences. The first type coincides with the sentence type É. Kiss has studied. Kálmán et al. (1986: 132) claim that "In every Hungarian corrective sentence there is what we refer to as an 'eradicating stress', i.e., a main stress that is not necessarily stronger than a normal stress, but which 'eradicates' all subsequent stresses, and thus, cannot be followed by any more main stresses". So, according to Kálmán et al. sentences with a single main stress may appear in Hungarian but they have a rather marked status. The unmarked order is represented by the sentence type which they call *neutral*. This type of sentence has no single prominent stress, and displays a 'level-prosody' pattern from an intonational point of view. Kálmán et al. claim that in corrective sentences all orders are possible, although there is a difference in interpretation associated with the various orders. Neutral sentences, on the other hand, allow only an SVO-order.[3] Compare the corrective sentences in (2) with their neutral counterpart in (3) (' indicates normal stress):

(2) There are a number of studies on the Focus-position in Hungarian including, among others, Farkas (1986), Hetzron (1966), Horvath (1976, 1986a), Hunyadi (1981b; 1981c), Jékel (1984), Kiefer (1967; 1981; 1986), É. Kiss (1981a; 1981b; 1981d; 1986b; 1987a), Kenesei (1985c; 1986b), Komlósy (1982a; 1986), and Szabolcsi (1980; 1981b; 1981c; 1983d).

(3) See for discussion of intonational and stress patterns in Hungarian: Kálmán (1985a; 1985b), Kálmán and Kornai (1985), Kenesei and Vogel (1986; 1987), É. Kiss (1987a), Komlósy (1986), Nádasdy (1985), Prószéky (1985), and Varga (1979; 1981a; 1981; 1983; 1985).

(2) a. 'Péter 'megvárta Marit a klubban
Peter perf-waited Mary-ACC the club-INESS
'Peter DID wait for Mary at the club.'
b. 'Péter 'Marit várta meg a klubban
'It is Mary that Peter waited for at the club.'
c. 'Péter 'várta meg Marit a klubban
'There has been an occasion when Peter waited for Mary at the club.'
d. 'Péter a 'klubban várta meg Marit
'It is at the club that Peter waited for Mary.'
e. 'Péter várta meg a klubban Marit
'It was Peter who waited for Mary at the club.'
f. 'Péter várta meg Marit a klubban
'It was Peter who waited for Mary at the club.'
(Kálmán et al. 1986: 131)

(3) a. 'Péter 'megvárta 'Marit a 'klubban
Peter perf-waited Mary-ACC the club-INESS
'Peter waited for Mary at the club.'
b. *'Péter 'Marit 'várta 'meg a 'klubban

I will consider this distinction to be observationally *adequate.* Therefore, following Kiefer (1967), Horvath (1981; 1986), and Kálmán et al. (1986), I will assume the following descriptive generalization on the *neutral* word order in Hungarian sentences:

(4) The neutral order is *SVO*

In the next section, I will argue that this order is derived from the underlying *SOV*-order by V-movement. Let us now consider the position of verbal modifiers in the sentential clause.

(*B*) Ackerman and Komlósy (cf. Ackerman 1984; 1987a, Ackerman and Komlósy 1983, and Komlósy 1985) observe another restriction on sentential word order in Hungarian. According to Ackerman and Komlósy, verbal *modifiers* must appear left-adjacent to the finite verb in neutral order. In such instances, the verbal modifier and the verb constitute a complex verb (cf. section 4.4.). The group of verbal modifiers which has this property is categorially rather heterogeneous and includes, among others, verbal prefixes[4] (cf. (5a)), determinerless complements of the verb (cf. (5b)), and predicative adjectives and nominals (cf. (5c)). Consider:

(5) a. Mari *be* dobta a labdát a tóba
Mary into threw-AGR3sg the ball-ACC the lake-ILL
'Mary threw the ball into the lake.'
b. János *fát* vágott az erdőben
John wood-ACC cut-AGR3sg the forest-INESS
'John was wood-cutting in the forest.'
(Ackerman 1984: 66)
c. *Beteg* lett
sick became
'He became sick'

These sentences support the following generalization on the position of verbal modifiers in their neutral order:

(4) See Kiefer (1982) for the role of verbal prefixes in the aspectual system of Hungarian.

(6) Verbal modifiers precede the finite verb in their neutral order

Note that the neutral order with a determinerless object is SOV (cf. (5b)). However, this is only an apparent violation of (4). In section 5.3.1., I will argue that such cases fall under the phenomenon of Noun-Incorporation which is conditioned by the absence of the definite or indefinite article. Let us now turn to a discussion of the syntax of Focus.

(C) Hungarian syntax is constrained by a fixed position for *Focus*-interpretation. With Kiefer (1967), among others, I will assume the following descriptive rule for this phenomenon:[5]

(7) The Focus-position is left-adjacent to the finite verb

That rule (7) is indeed operative in Hungarian may be observed from the *Inversion* between the verb and the verbal modifier when a constituent, apart from verbal modifiers themselves, is focussed. Focussed NPs and verbal modifiers are in complementary distribution.[6] Compare the minimal pair (3a) versus (2b), here repeated as (8a) and (8b):

(8) a. 'Péter *'meg* várta 'Marit a 'klubban
Peter perf-waited-AGR3sg Mary-ACC the club-INESS
'Peter waited for Mary at the club.'
b. 'Péter 'Marit várta *meg* a klubban
'It is Mary that Peter waited for at the club.'

In the neutral (8a), the verbal modifier, the prefix *meg* 'perfectivity marker', precedes the verb, whereas in (8b) in which the accusative NP is focussed, it must be postponed.

Other NPs with quantificational content trigger also Inversion. Wh-phrases in Hungarian occupy the Focus-position, because they must be left-adjacent to the finite verb. As a consequence, with Wh-questions the verbal modifier has to be postponed:

(9) a. *Ki *meg* látta Marit?
who perf- saw-AGR3sg Mary-ACC
('Who did notice Mary.')
b. Ki látta *meg* Marit?
'Who did notice Mary.'

These sentences support the following descriptive generalization on the position of *Wh-phrases:*

(10) Wh-phrases appear in the Focus-position

É. Kiss (1981b: 189) lists some other NPs with quantificational content which have to appear in Focus-position obligatorily involving, among others, constituents

(5) This descriptive statement is incorporated into a formal approach by É. Kiss (1981) and Horvath (1986). É Kiss puts this restriction into her phrase structure rules of Hungarian, while Horvath assumes that each Hungarian verb is associated with a Focus-feature which is assigned to the maximal projection to the left of the verb under strict local government.

(6) É. Kiss (1981b) refers to the category of verbal modifiers as 'reduced' complements. According to É. Kiss, reduced complements are in Focus. On the other hand, Ackerman and Komlósy (1983) point out, correctly in my view, that although verbal modifiers and focussed constituents are in complementary distribution, this does not imply that the verbal modifiers occupy the Focus-position in their neutral order.

modified by a negative particle, or by *csak* 'only'. Consequently, they also trigger Inversion between a finite verb and a verbal modifier. Compare:

(11) a. *Nem János meg látta Marit
not John perf- saw-AGR3sg Mary-ACC
('Not John did notice Mary.')

b. Nem János látta *meg* Marit?
'Not John did notice Mary.'

(12) a. *Csak János *meg* látta Marit
only John perf- saw-AGR3sg Mary-ACC
('Only John did notice Mary.')

b. Csak János látta *meg* Marit?
'Only John did notice Mary.'

Hence, Inversion is captured by the following generalization:

(13) Focussing triggers Inversion between the finite verb and its verbal modifier

Let us consider now the position of quantified expressions in Hungarian.

(*D*) Quantifiers prefer a position to the *left* of the finite verb (see, for example Hunyadi 1981a, among others).[7] So, not only focussed NPs have to be to the left of the finite verb but in fact any constituent with a quantificational content. Kenesei (1986) regards these phenomena as subcases of the same restriction on word order in Hungarian. According to Kenesei, elements with a quantificational content, such as negated NPs, inherent quantifiers, Wh-phrases, focussed NPs, and so on, occupy a "field" to the left of the verb. Kenesei distinguishes the following four fields in the Hungarian sentence:

(14)

Initial Field	*Quantifier Field*	*Verb*	*Postverbal Field*
non-operators (i.e. 'Topics', existential *Q*, downgraded universal *Q*	*even*/*no*-phrases > negation > universal *Q* > only-phrase/ Wh-phrase/ Focus		non-operators, *no*-phrases/ universal *Q*, existential *Q*, *even*-phrase

(Kenesei 1986: 148)

In (14), the slant lines stand for a disjunctive relationship and the 'greater than' ('>') sign for a strict left-to-right order. Kenesei claims that the order of constituents is rather free in the Initial Field and the Postverbal Field but that it displays a linear ordering in the Quantifier Field. Furthermore, he observes that scope-interpretation is a function of linear order. Kenesei postulates the following ad-hoc descriptive device to capture scope-readings:[8]

(15) Given quantifiers Q_1 and Q_2 where Q_1 precedes Q_2, Q_1 has scope over Q_2

(7) There is a lively discussion in this area of Hungarian grammar. The outcome of this debate might have important consequences for theories on the relation between syntax and semantics. Compare, among others, Bánréti (1982), Hunyadi (1981a; 1981b; 1984; 1985; 1986a; 1986b; 1987), Kenesei (1985b; 1985c; 1986b; 1987; to appear) Kiefer (1981; 1986), É. Kiss (1986b; 1987a), Marácz (1985a; 1986a), Ruzsa (1986), Szabolcsi (1980; 1981b; 1981c; 1983d; 1986a; 1986b), and Varga (1980).

(8) (14) does not cover several scope-readings. For example, a stressed universal quantifier in the Postverbal Field may have scope over a quantified expression in the Quantifier Field (' indicates stress). Compare:

(i) Csak Jánost szereti 'mindenki
only John-ACC loves everyone
'For every x, only for y=John, x loves y'
*'Only for y=John, for every x, x loves y'

Observe that the scope-readings in (16) are covered by (15):

(16) a. Mindenki csak Máriát szereti
everyone only Mary-ACC loves
'Everyone is such that he loves only Mary.'
*'Only Mary is such that everyone loves her.'
b. Csak Máriát szereti mindenki
*'Everyone is such that he loves only Mary.'
'Only Mary is such that everyone loves her.'

Kenesei notes some further restrictions in the Quantifier Field. For example, Wh-phrases cannot be preceded by any NP with quantificational content other than another Wh-phrase. This is illustrated in the following pair:

(17) a. *Mit/*valakit/*mindenkit/*egy férfit/*csak téged* ki
what-ACC/someone-ACC/everyone-ACC/a man-ACC/only you-ACC who
látott?
saw-AGR3sg
'Who saw what?'
b. Ki *mit* látott *valakit/mindenkit/egy férfit/csak téged?*
'Who saw what/someone/everyone/a man/only you?'
(Kenesei 1986: 153)

In order to make this descriptive generalization more explicit, Kenesei (1986: 153) formulates schemes which have the effect of restricting rule (15):

(18) *[S NP_1[-Wh]...NP_2[+Wh]] where NP_2 is in the scope of NP_1

Summarizing, quantifiers in Hungarian appear preferably "stacked" to the left of the finite verb (cf. (14)). Their scope-interpretation is determined by the linear order in which they appear in the sentence (cf. (15)). This may further be restricted by the content of the quantifiers (cf. (18)). It goes without saying that both the position and the interpretation of quantifiers heavily constrain the freedom of word order.

Let us consider now the word order in maximal projections other than the sentential clause.

(*E*) In general, maximal projections other than the sentential clause are *head-final.* Within a single maximal projection complements precede their heads. Therefore, we may formulate the following descriptive generalization on the relative order of complement and head:

(19) Endocentric categories are head-final

Observe, for example, that an NP, a PP, and a participle construction, which is an NP in Hungarian, have their head on the right periphery:[9]

Furthermore, scope is not determined by word order with the existential quantifier *valaki* 'someone':

(ii) *Valakit mindenki szeret*
someone-ACC everyone loves
'For every x, for some y, x loves y'
'For some y, for every x, x loves y'

It is easy to see that the first reading is not predicted by rule (15). From these examples, I conclude that (15) can be overridden by phonological and lexical factors. This implies also that it does not give a complete picture in itself of scope-assignment in Hungarian (cf. Hunyadi 1981a and Kenesei 1986 for suggestions).

(9) Studies on the NP include, among others, Dezső (1967; 1971; 1982a), Gaál (1978), Kenesei (1985e), Kornai (1985), Szabolcsi (1981a), and Tompa (1968). For discussion of the PP compare Marácz (1983; 1984; 1985c; 1986c), Papp (1963), and Sebestyén (1965). In chapter seven, I will return to the structure of NPs and PPs in more detail.

(20) a. A piros *ház*
the red house
'The red house.'

b. A ház *mögött*
the house behind
'Behind the house.'

c. A sarkon álló *ház*
the corner-SUPER stand-pres.part. house
'The house which stands at the corner.'

Observe that (19) holds only for endocentric categories which are a projection of the expansion of their heads. Furthermore, from (19) it follows that maximal projections are *left-branching.*

Let us consider another example of an endocentric category in Hungarian, the possessive NP.

Szabolcsi (1981a) has observed that this construction displays two variants. A variant in which the possessor NP is marked nominatively, and a variant in which the possessor NP appears with the dative case. In both constructions, the possessor NP precedes the noun-possessed, the head of the possessive NP. The noun-possessed bears an person-number agreement (glossed as npAGR).[10] Compare:

(21) a. A fiú háza
the boy house-npAGR3sg
'The house of the boy'

b. A fiúnak a háza
the boy-DAT the house-npAGR3sg
'The house of the boy'

Szabolcsi (1981a) has observed some further syntactic differences between these variants.

(*i*) The definite article a(*z*) invariably precedes the nominative possessor NP (cf. (21a)), whereas it invariably follows the dative possessor NP (cf. (21b)).

(*ii*) The nominative possessor may not be separated from the head noun. The dative possessor, on the other hand, may scramble freely around in the sentence. Consider:

(22) a. *A *fiú* leégett [*t* háza]
the boy down-burned house-npAGR3sg
'The house of the boy burned down.'

b. A *fiúnak* égett le [*t* a háza]
the boy-DAT burned down the house-npAGR3sg
'The house of the boy burned down.'

From this minimal pair, Szabolcsi concludes that the dative possessor NP, unlike the nominative possessor NP, does not have to be in construction with its noun-possessed.

(*iii*) Wh-phrases may only occur as a dative possessor NP:

(23) a. *A *ki* vendége
the who guest-npAGR3sg

b. *Kinek* a vendége
who-DAT the guest-npAGR3sg
'Whose guest'

The question arises what happens when a right-branching category is embedded in a left-branching endocentric category? This can only happen if Hungarian had

(10) See for studies of the Hungarian possessive NP, among others, Biermann (1985), Gaál (1978), De Groot (1983b), Kenesei (1985e), Kornai (1984; 1985), Mel'cuk (1973), and Szabolcsi (1981a; 1984; 1986c; 1986d; 1986e; 1986g; 1987c).

right-branching exocentric categories. Kenesei (1984) argues that relative clauses are such. Compare:

(24) a. [NP NP [CP ...]] b. [NP A fiú [CP aki a sarkon áll]]
the boy who the corner-SUPER stands
'The boy who is standing at the corner.'

Consider now the output of embedding a relative clause in a possessive NP (cf. (25)), or PP (cf. (26)):

(25) a. *[[NP A fiú [CP aki a sarkon áll]] köpenye]
the boy who the corner-SUPER stands cloak-npAGR3sg
b. [[NP A sarkon álló fiú] köpenye]
the corner-SUPER stand-pres part boy cloak-npAGR3sg
'The cloak of the boy who was standing on the corner.'

(26) a. *[PP [NP A ház [CP amely a sarkon áll]] mögött]
the house which the corner-SUPER stands behind
b. [PP [NP A sarkon álló ház] mögött]
the corner-SUPER stand-part.pres. house behind
'Behind the house on the corner'

These sentences demonstrate that in left-branching endocentric categories no right-branching categories may appear. In the grammatical variants, the relative clause has been transformed into a left-branching category by an adjectivizing strategy. This category is headed by the present participle which modifies the complement of the noun-possessed or postposition. With respect to the possessive NP, there exists another strategy to save configuration (25a), namely, by marking the possessor NP with dative case:

(27) [NP A fiú*nak* [CP aki a sarkon állt]] véres volt [*t* a köpenye]
the boy-DAT who the corner-SUPER stood bloody was the cloak-npAGR3sg
'The cloak of the boy who was standing at the corner was bloody.'

Recall, however, that a dative possessor NP may be scrambled out of its possessive NP yielding a discontinuous constituent. This suggests that generalization (19) holds if and only if the head and its complement are in construction. Therefore, a dative possessor NP, which is separated from its noun-possessed, may head a right-branching structure.

Recapitulating, in this section I have discussed the following descriptive generalizations on word order in Hungarian:

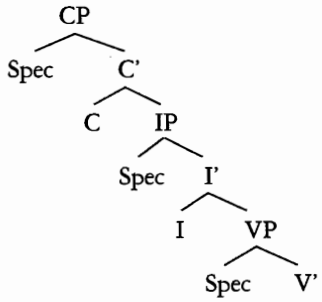
(28) a. The neutral order is *SVO* (cf. (4))
b. Verbal modifiers precede the finite verb in their neutral order (cf. (6))
c. The Focus-position is left-adjacent to the finite verb (cf. (7))
d. Wh-phrases appear in the Focus-position (cf. (10))
e. Focussing triggers Inversion between the finite verb and its verbal modifier (cf. (13))
f. Quantifiers are stacked preverbally (cf. (14))
g. Given quantifiers Q_1 and Q_2 where Q_1 precedes Q_2, Q_1 has scope over Q_2 (cf. (15))
h. Endocentric categories are head-final (cf. (19))

In the next section, I will rely heavily on these descriptive generalizations in the development of a theory for the phrase structure of Hungarian.

2.2. Theory

In this section, I will present a theory for the *phrase structure* of Hungarian. Following Chomsky (1986a), I will assume that the categorial component of the grammar universally generates the following phrase structure:

(1)



Let us consider then how the descriptive generalizations of the preceding section fit into (1).

2.2.1. Hungarian is an *SOV-language*

In chapter five, I will argue on the basis of various subject-object asymmetries that Hungarian has a VP-node. The next question to anwer is what the basic order of this category is.

Recall that endocentric categories in Hungarian are head-final (cf. 2.1.(28h)) and that these categories may not contain right-branching substructures. This implies that the general directionality of branching in Hungarian is leftward in its endocentric projection. Therefore, I propose the following uniformity condition on the directionality of branching of X'-categories:

(1) *Uniformity Condition on Branchingness of X'-Categories*
Endocentric categories are left-branching in Hungarian

This principle is due to a core principle of the X'-component which generates only right- or left-branching categories in a particular language (cf. Stowell 1981).

The following phenomena also illustrate that (1) is operative in Hungarian.

(*I*) Apart from NP and PP, the VP is head-final as well, although in contemporary Hungarian (Hungarian is historically an SOV-language (cf. Bárczi et al. 1978) the OV-order does not surface in finite sentences for reasons having to do with V-movement (see the next section for discussion). However, constructions with non-finite verbs, like participle constructions, are unambiguously head-final (cf. 2.1.(25b) and 2.1.(26b)).

(*II*) Left-branching categories may not contain right-branching substructures, as the ungrammaticality of 2.1.(25a) and 2.1.(26a) indicates. Hungarian employs seve-

ral adjectivizing strategies to circumvent this type of embedding. For example, by inserting 'dummy' verbal participles (cf. the case of embedded relative clauses discussed in 2.1.(25a) and 2.1.(26a)) or by adjectivization with the suffix *-i* (cf. also Laczkó 1985 and section 7.2.1.).

Nominalization is also an instance of the former strategy. The verbs *átkel* 'cross over' and *tartoz* 'belong to' may be nominalized by suffixing of *-ás/és* (NOMI). *Átkel* subcategorizes for an NP with a lexical superessive case (cf. (2a)) and *tartoz* subcategorizes for a lexical allative case (cf. (3a)). Nominalizations with *-ás/és* are instances of passivization (cf. chapter three):

(2) a. NP átkel a hídon
NP cross-AGR3sg the bridge-SUPER
'NP crosses over the bridge.'
b. [NP [NP az átkelés] [NP a hídon]]
the cross-NOMI the bridge-SUPER
'The crossing over the bridge'

(3) a. NP tartozik a csoporthoz
NP belong-AGR3sg the group-ALL
'NP belongs to the group.'
b. [NP [NP a tartozás] [NP a csoporthoz]]
the belong-NOMI the group-ALL
'The belonging to the group'

The (b)-phrases demonstrate that a nominalized verb may take an NP to its right. This NP is case-marked similarly as the NP-complement of the unmodified alternant in the (a)-phrases.

The following examples show that the insertion of adjectivers, like the verbal participles *való* 'being' of the verb *van* 'be' and *történő* 'happening' of the verb *történik* 'happen', may transform the right-branching structures in the (b)-phrases into left-branching structures. (*Való* is a stative present participle and *történő* is a dynamic present participle):

(4) a. [NP a hídon *történő* átkelés]
the bridge-SUPER happen-part cross-NOMI
'The crossing over the bridge'
b. [NP a csoporthoz *való* tartozás]
the group-ALL be-part belong-NOMI
'The belonging to the group'

Another strategy to create left-branching structures is by adjectivization with the suffix *-i* (adj). Consider the following phrases:

(5) a. [NP [NP a lány] [NP Budapestről]]
the girl Budapest-DELAT
'The girl from Budapest'
b. [NP [NP a folyó] [PP a híd alatt]]
the river the bridge under
'The river under the bridge'
c. [NP [NP János kémkedése] [PP a főnök után]]
John spy-NOMI-npAGR3sg the boss after
'John's spying upon the boss'

In the above phrases, a (possessive) NP takes an NP (cf. (5a)) or a PP (cf. (5b) and (5c)) to its right. These phrases may be turned into left-branching categories by suffixing the latters with the adjectivizer *-i* :

(6) a. [NP a [NP budapest]$_i$ lány]]
the Budapest-adj girl
'The girl from Budapest'

b. [NP a [PP híd alatt]$_i$ folyó]]
the bridge under-adj river
'The river under the bridge'

c. [NP Jánosnak [PP a főnök után]$_i$ kémkedése]
John-DAT the boss after-adj spying-NOMI-3npAGR
'John's spying upon the boss'

Laczkó (1985) reports that the types of adjectivization in (4) and (6) are quite common and that they are preferred over their right-branching counterparts.

The cross-category generalization in (1) has far-reaching implications for the phrase structure of Hungarian. As I pointed out above, the VP is underlyingly *OV*. This implies that Hungarian is an *SOV*-language. Furthermore, the VP cannot contain right-branching substructures. Therefore, Horvath's (1981, 1.6.3.) argument for a right-branching V" based on Emonds' restriction on surface recursion, must be rejected on conceptual grounds (cf. Ackerman 1984). Let us now consider how the neutral *SOV*-order is covered (cf. 2.1.(28a)).

2.2.2. *V-movement and the IP-parameter*

I will assume that the neutral *SVO*-order is derived from the underlying *SOV*-order, analogously with the rule which yields the V-second effect in Germanic languages. In these languages, the verb is in final position in embedded clauses, but it is in the second position in root clauses.

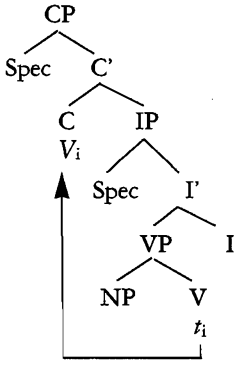
Let us consider an example of V-second in Dutch:

(1) a. Jan dacht dat Peter hem gisteren *opbelde*
John thought that Peter him yesterday up-phoned

b. Peter *belde* hem gisteren *op*
Peter phoned him yesterday up

In (1a), the verb *opbellen* 'to phone up' is in its base-generated position in the embedded sentence. In (1b), on the other hand, the finite verb appears in the second position of the root clause and it leaves its particle stranded in base-generated position. It has been argued that V-second is derived by V-movement in root clauses (cf. Koster 1975, Thiersch 1978, Haider and Prinzhorn 1986, and Platzack 1982, among others). V-second yields the following configuration:

(2) *V-second*



Note that the *C* position acts as the landing site for the moved finite verb. The question is then what triggers V-movement.

Koster (1986) argues that this has to do with the status of C in Germanic languages. According to Koster, the governors I and C have different properties from ordinary lexical governors. Henceforth, I will refer to them as *functional* governors.

Normally, lexical governors, like V, N, P, or A, determine a syntactic minimal maximal *domain,* i.e. VP, NP, PP and AP, and control a *Case-position.* Functional governors do not always display these properties.

The projections of these governors, CP and IP, are auxiliary projections to VP. This entails, among other things, that the local domain of categories governed by V is not necessarily VP but may be CP, for example. Lexical governors assign argument status to the NPs they govern, unlike the functional governors. Neither C, nor I assigns a θ-role to any NP. I is usually assumed to be associated with nominative Case, but C does not even need to assign Case.

Because of this dichotomy between lexical governors on the one hand and functional governors on the other hand, Koster argues that the CP- and IP-projections should not be treated on a par with the projections of lexical governors. Lexical governors are always *strong* in the sense that they determine a projection, and may control a Case-position. However, the 'strength' of functional governors may vary. C or I can be strong or weak. With Koster (1986), I will hypothesize that the strength of governors is defined as in (3a), and that strong governors have the syntactic properties in (3b):

(3) a. A governor is strong if it can be *lexically* filled, otherwise it is weak
b. A strong governor determines a *projection,* and controls a *Case-position*

Furthermore, I will hypothesize that at least one of the functional governors must be strong. This is probably due to the requirement that a clause must be complete functionally. Outside the VP, there must be a position available for the external argument, the subject. This can only be guaranteed if either CP or IP is present. This implies then the following:

(4) Either *C* or *I* is a *strong* governor

So in order to determine the strength of a governor, we must check whether there is independent lexical material available to fill the position of that governor. Let us consider the strength of the functional governors in the Germanic languages.[11]

In all Germanic languages, C is strong because these languages possess lexical complementizers. As a consequence, all these languages realize a CP-projection, at least in embedded clauses. What, on the other hand, is the strength of I? It is generally assumed that I is lexically filled in English by auxiliary verbs, like *do,* modal verbs, such as *can* or *may* (cf. Steele 1981).[12] Hence, it is strong in English. If it is strong, I creates its own domain, namely IP, and it assigns nominative Case to the

(11) In this chapter, I will restrict myself to Dutch, English, Frisian, and German. See for a discussion of V-movement in Scandinavian Koster (1986) and Platzack (1982; 1987), among others.

(12) Koster (1986) observes two apparent exceptions to the claim that I is always lexically filled in English.

subject. In the other Germanic languages, however, there are no independent lexical items for the I-position available. Hence, I is weak. This yields the following parameter:

(5) *IP-parameter*
a. I is strong in English; b. I is weak in Dutch, Frisian, and German

Let us consider some implications of the IP-parameter for the syntax of these languages.

Both C and I are weak in the root clauses of the other Germanic languages, because they remain lexically unfilled. Note that this state of affairs violates principle (4). How do these languages escape this conflict?

Following Koster (1986), I will assume that movement of V to C turns C into a strong governor, for C gets lexically filled by the moved verb. This yields the V-second effect. Hence, there seems to be a tight relation between V-movement and the strength of the governor in which it lands. V-movement is triggered by a strong governor. The question, then, is why V-to-C movement does not occur in English.

V-to-C movement must proceed stepwise, as required by Chomsky's (1986a) *Head Movement Constraint* which I will define as follows:

(6) *Head Movement Constraint* (HMC): An X^0 may move into a Y^0 that governs it

Because of (6), V must first move to I before it can reach C. In English, I cannot function as an extraction-site for V-to-C movement, since I is always filled lexically. As a consequence, C remains unfilled in English root clauses.[13] Note, however, that this does not violate (4). So V-to-C movement applies only under the following conditions:

(7) V-to-C movement applies if and only if C is strong and I is weak

Let us now determine the strength of the functional governors in Hungarian. In Hungarian, there are no independent lexical items, such as auxiliaries or modals in English, to fill I. Hence, I is weak. C, on the other hand, is strong, for Hungarian possesses lexical complementizers, like *hogy* 'that'. Hence, we derive the following:

(8) a. *C* is a strong governor, and; b. *I* is a weak governor in Hungarian

---

(i) Sentences without I-fillers, like (ia):
(i) a. They beat horses
b. They do not beat horses
Koster argues, however, that in the D-structure representation of (ia), I is filled with *do*, similar to its negative counterpart (ib). Do is, however, deleted at S-structure in (ia).

(ii) C must sometimes be filled by the movement of I:
(ii) a. *[$_{CP}$ What [$_{IP}$ he has done?]]
b. [$_{CP}$ What has$_j$ [$_{IP}$ he t$_j$ done?]]
These sentences show that Wh-movement to [Spec, CP] triggers I-to-C movement. This is probably due to the requirement that a position in a projection is only available if the head of this projection is lexical or a trace of a lexical item.

(13) Except for the case of Subject-Aux Inversion. With this phenomenon, C is filled by the movement of I to C. See note 12(ii) and section 5.4.3.1. for discussion of I-to-C movement in English root clauses.

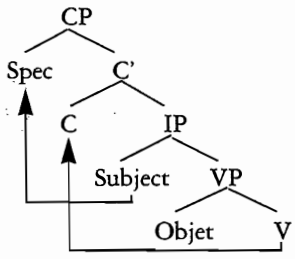
Observe that the functional governors have precisely the same strength in Hungarian as in Dutch, Frisian, or German. If V-to-C movement in these languages is triggered by a strong C and not blocked by a weak I (cf. (7)), then it follows that V-to-C movement applies in Hungarian as well. Hence, this yields the following hypothesis:

(9) *V-movement Hypothesis for Hungarian:* *V* moves to *C* in finite sentences

Below, I will argue that V-movement is 'generalized' in Hungarian. It does not only apply in root clauses but also in embedded clauses. This is allowed because, as I will attempt to demonstrate, CP is recursive within CP. Let us first consider some empirical evidence for (9).

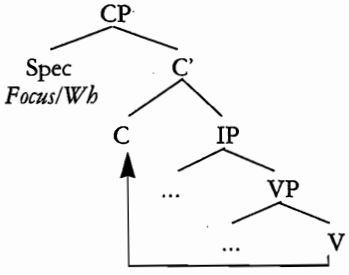
(*I*) The fact that the neutral order in Hungarian is *SVO* indicates that (9) is operative. The application of V-movement to the underlying *SOV*-order and movement of the subject to the [Spec, CP] position yields an *SVO*-order (cf. 2.1.(28a)). This is depicted in the following diagram:

(10)



(*II*) If V-to-C movement results in a V-second effect in Germanic languages, then we expect such an effect in Hungarian as well. A property of Hungarian which resembles V-second is the adjacency requirement on the Focus-position (cf. 2.1.(28c)). Recall that Focus must be left-adjacent to the finite verb. Let us interpret this requirement as the Hungarian manifestation of V-second. Hence, a sentence with a filled Focus-position has the following configuration:

(11)



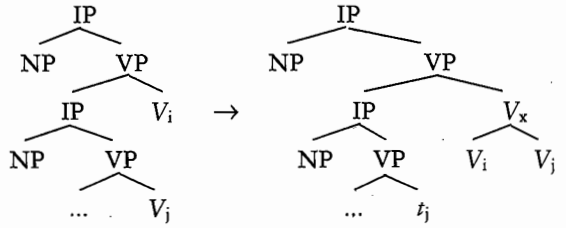
Note from (11) that under this approach Focus equals the [Spec, CP] position. A wellcome consequence of this is that Wh-phrases in Hungarian occupy also the [Spec, CP] position (cf. 2.1.(28d)), similarly to Wh-phrases in Germanic.

A concomitant of V-second is that particles of particle-verb combinations must remain stranded in their base-generated position (cf. (1)). We expect then that with the V-second effect in Hungarian, i.e. Focussing, verbal particles may not be moved

along with the verb. This turns out to be the case. Recall that Inversion between the finite verb and its verbal modifier, including particles, prefixes, and so on, is obligatory under Focussing (cf. 2.1.(28e)). Hence, "Inversion" is due to movement of the verb leaving its verbal modifier stranded.

(*III*) Apart from V-second, Dutch has another instantiation of V-movement, namely, V-*raising* (cf. Evers 1975). This type of V-movement adjoins an infinitival clause without a complementizer that is base-generated on the left-hand side of the verb of the next higher clause to the right of this verb. This yields the following configuration:

(12) *V-raising*



The following sentences exemplify some instances of *V-raising:*

(13) a. *Ik geloof [CP dat [IP Jan [IP Nederlands te leren]] begint]
I believe that John Dutch to learn begins
b. Ik geloof [CP dat [IP Jan [IP Nederlands t_j]] begint *te leren*$_j$]
I believe that John Dutch begins to learn

Van Riemsdijk and Williams (1986) report that the local character of V-raising manifests itself with (*A*) *adjacency effects* and (*B*) *restructuring phenomena*. Let us first discuss (*A*).

(*Ai*) Consider a V-raising construction with a verb combining with a particle and selecting a left-hand infinitival complement:

(14) ...[IP ... V] Particle V...

An example of this configuration is (15):

(15) *Ik geloof [CP dat [IP Jan [IP Nederlands t_j]] aanvangt *te leren*$_j$]
I believe that John Dutch starts to learn

V-raising is blocked in configuration (14). The reason for this is that the particle *aan* of the verb *aanvangen* 'to start' intervenes between the higher verb and the V-raised verb. Hence, the adjacency requirement on V-raising is violated.

(*Aii*) Certain PPs in Dutch, may optionally be extraposed to the end of the clause in which they appear. Compare:

(16) a. Ik geloof dat Jan probeert [IP de schuur [PP *met een spraydoos*] groen te
I believe that John tries the barn with a a spray can green to
schilderen]
paint
'I believe that John is trying to paint the barn green with a spray can.'
b. Ik geloof dat Jan probeert [IP de schuur groen te schilderen [PP *met een spraydoos*]]

When this PP-extraposition occurs in a left-hand complement of a V-raising verb, a structure like (17) occurs:

(17) ...[IP... V PP] V ...

Note that V-raising cannot apply to this structure:

(18) *Ik geloof dat [IP Jan [IP een schuur groen t_j [PP *met een spraydoos*]]] wil *schilderen*$_j$
I believe that John a barn green with a spray can wants to paint

The ungrammaticality of (18) demonstrates that V-raising is sensitive to an adjacency requirement.

(*Aiii*) If V-raising is conditioned by an adjacency requirement, then we expect that the mirror-image of the D-structure order is derived when several left-hand side infinitival complements are embedded within each other. The following pair shows that this turns out to be the case:

(19) a. Ik geloof dat Jan [VP [IP *PRO* [VP [IP *PRO* [VP [IP *PRO* [VP dit boek lezen$_1$]] leren$_2$]] proberen$_3$]] zal$_4$]
I believe that John this book read learn try will
'I believe that John will try to learn to read this book.'
b. Ik geloof dat [IP Jan dit boek] zal$_4$ proberen$_3$ te leren$_2$ lezen$_1$.

(*B*) Let us now consider the restructuring phenomena. So, in V-raising constructions, a V-head of a complement clause is adjoined to the V of the next higher clause. As a result, a complement clause without a complementizer may become transparent with respect to government. For example, Koster (1987: chapter three) discusses the following restructuring effects involving, among others, NP-raising, Exceptional Case Marking, Obligatory Control, Transparency, R-movement, Adverbial Scope, and Clitic Climbing. Let us discuss, for instance, Clitic Climbing.

Koster reports that clitics, like Dutch *het* 'it', can be moved across subjects in V-raising complements:[14]

(20) Ik denk dat hij *het*$_i$ [IP Peter [IP t_i t_j] hoorde zingen$_j$]
I think that he it Peter heard sing
'I think that he heard Peter sing it.'

Note that *het* has been moved from its object position in the complement clause across the subject constituent *Peter* of the embedded clause. According to Koster, this is a striking fact, because normally *het* cannot be moved across a subject. This kind of "clitic climbing" is possible only from V-raising complements. It is never possible to move *het* out of an extraposed *om*-complement. Hence, the ungrammaticality of (21b):

(21) a. Ik denk dat Peter probeerde [CP om [IP *het* aan Mary te geven]]
I think that Peter tried COMP it to Mary to give
'I think that Peter tried to give it to Mary.'
b. *Ik denk dat Peter *het* probeerde [CP om [IP *t* aan Mary te geven]]
I think that Peter it tried COMP to Mary to give

(14) Following Koster and May (1982), I will assume that infinitival phrases are clauses and that tensed and infinitival clauses share the same phrase structure. Hence, they are IPs.

It has been observed in the literature (cf. Evers 1982 and de Haan 1982, among others) that V-raising appears in languages with a V-second effect. Moreover, Evers (1982) even argues that these types of verb movements are different instantiations of the same principle. In any case, we therefore may postulate the following implication:

(22) If a language *X* displays V-raising, then *X* also displays V-movement

From this it follows that the occurrence of V-raising in a particular language provides an *indirect* argument for V-movement in that language. Let us consider then V-raising appears in Hungarian.

Kálmán et al. (1986) have observed that Hungarian has two groups of verbs which may select infinitival complements, namely *auxiliary verbs* and *main verbs*. The former group includes, among others, *akar* 'want', *bír* 'can', *fog* 'will', *kell* 'have to' (impers.), *kezd* 'begin', *kíván* 'wish to', *lehet* 'it is possible to; one can' (impers.), *mer* 'dare', *méltóztatik* 'be pleased to; one can' (impers.), *próbál* 'try to', *szabad* + copula 'it is permitted to' (impers.), *szándékozik* 'wish to' (no definiteness agreement), *szeretne* 'would like to', *szokott* 'used', *tetszik* 'be pleased to' (auxiliary of polite verb forms, impers.), and *tud* 'can'. The group of main verbs includes, among others, *utál* 'hate', *imád* 'adore', *elfelejt* 'forget', *szeret* 'like to', *enged* 'allow', *megy* 'go', and *vél* 'believe'.

Consider the following examples:

(23) a. János [$_{IP}$ úszni] akart (auxiliary)
John swim-INFI wanted-AGR3sg
'John wanted to swim.'

b. János imádott [$_{IP}$ sétálni Marival] (main verb)
John loved-AGR3sg walk-INFI Mary-INSTR
'John loved to walk with Mary.'

c. Péter [$_{IP}$ játszani] ment (main verb)
Peter play-INFI went-AGR3sg
'Peter went to play.'

Sentence (23a) shows that in neutral order an infinitival complement occurs on the left-hand side of the auxiliary verbs. Sentence (23b) demonstrates, however, that the infinitival complements occur on the right-hand side of main verbs in their neutral order, except with the verbs *megy* 'go' (cf. (23c)) and *vél* 'believe'.

From the examples in (23) V-raising cannot be proved. One could argue that the finite verbs in (23a) and (23c) remain, for some reason, in their base-generated order, and the finite verb in (23b) skips over its infinitival complement by V-movement (cf. (8)). Note therefore the following sentences:

(24) a. János [$_{IP}$ el t_j] akart *úszni*$_j$
John away wanted-AGR3sg swim-INFI
'John wanted to swim away.'

b. János imádott [$_{IP}$ elsétálni Marival]
John loved-AGR3sg away-walk-INFI Mary-INSTR
'John loved to walk away with Mary.'

c. Péter [$_{IP}$ t_i beiratkozni] ment *az iskolába*$_i$
Peter in-register-INFI went-AGR3sg the school-ILL
'Peter went to register with the school.'

Auxiliaries induce "Aux-splitting" when they select an infinitival complement which is itself modified by a verbal modifier. The auxiliary *akar* must obligatorily appear between the prefix *el* and the infinitive *úszni* of the particle-infinitive combination *elúszni* 'to swim away' (cf. (24a)). Main verbs, on the other hand, do not trigger Aux-splitting. Let us concentrate on the infinitive constructions with auxiliary verbs.

Aux-splitting cannot be derived by movement of the finite verb into the infinitival complement, because this would violate the c-command condition on traces. If this option is ruled out, then the only possibility to derive Aux-splitting is by *V-raising* as indicated in (24a). V-raising of the infinitive leaves the particle stranded in its base-generated position. Hence, Hungarian displays V-raising.

Above, I noted that V-raising has two sorts of diagnostics. It exhibits locality and restructuring effects. In section 5.3.2., I will argue that restructuring phenomena with V-raising appear in Hungarian as well. These phenomena involve, among others, some auxiliaries displaying person-number agreement with the object NP of their infinitival complement, and obligatory subject control.

Adjacency effects are much harder to prove with V-raising in Hungarian, because it allows scrambling. For example, a sentence adverb, like *tegnap* 'yesterday', may intervene between the auxiliary verb and a V-raised infinitive. Compare the counterpart of (24a):

(25) János [$_{IP}$ el t_j] akart *tegnap* úszni$_j$
John away wanted-AGR3sg yesterday swim-INFI
'John wanted to swim away yesterday.'

Locality effects, however, appear with the stacking of V-raised infinitives. Recall that the order in which V-raised infinitives are attached to the higer verb is precisely the opposite of the D-structure order (cf. (19)). Kenesei (1985c) has observed that this also appears in Hungarian. Consider the following sentences:

(26) a. János [$_{VP}$ [$_{IP}$ *PRO* [$_{VP}$ [$_{IP}$ *PRO* [$_{VP}$ a biciklit szétszedni$_1$]] tudni$_2$]] fogja$_3$]
John the bike-ACC apart-take-INF] can-INF] will-AGR3sg
'John will be able to take apart the bike.'
b. János [$_{VP}$ [$_{IP}$ szét] fogja$_3$ tudni$_2$ szedni$_1$] a biciklit
c. *János [$_{VP}$ [$_{IP}$ szét] fogja$_3$ szedni$_1$ tudni$_2$] a biciklit

In (26a), V-raising obligatorily applies yielding Aux-splitting. The deepest embedded infinitive may not occur in the derived structure between the auxiliary and the infinitive which is directly embedded under this auxiliary at D-structure (cf. (26c)). Only the reversed order is grammatical (cf. (26b)).

This locality effect is demonstrated even more persuasively in (27). Note that in these sentences the embedded infinitives are both prefixed. The infinitive *próbálni* 'to try' is prefixed with the perfectivity marker *meg* and the infinitive *úszni* 'to swim' is prefixed with *el* 'away'. Compare:

(27) a. János [$_{VP}$ [$_{IP}$ *PRO* [$_{VP}$ [$_{IP}$ *PRO* [$_{VP}$ a parttól elúszni$_1$]]
John the beach-ALL away-swim-INFI
megpróbálni$_2$]] akart$_3$]
pref-try-INFI wanted-AGR3sg
'John wanted to try to swim away from the beach.'

b. *János [VP [IP el] akart$_3$ megpróbálni$_2$ úszni$_1$] a parttól
c. *János [VP [IP el] akart$_3$ úszni$_1$ megpróbálni$_2$] a parttól
d. *János [VP [IP meg] akart$_3$ elúszni$_1$ próbálni$_2$] a parttól
e. János [VP [IP meg] akart$_3$ próbálni$_2$ elúszni$_1$] a parttól

Structure (27a) represents the underlying order of this paradigm. V-raising has to apply, because these infinitives have a prefix. Note now that only the prefix of the deepest embedded infinite may remain stranded and that the derived order must be the mirror-image of the D-structure order. Hence, only (27e) yields a grammatical result.

In conclusion, the locality effects in the paradigms of (26) and (27) strongly suggest that V-raising applies in Hungarian. If that is correct and implication (22) holds, then we provided an argument for the existence of V-movement.

So far I did not discuss generalization 2.1.(28b) which states that verbal modifiers precede the finite verb in their neutral order. Let us consider how this fits into the system outlined above.

With V-movement in Dutch, the particle remains obligatorily stranded in its base-generated position. This is illustrated by the following pair:

(28) a. Peter beldej hem gisteren *op* t_j b. *Peter *op*belde$_j$ hem gisteren t_j
Peter phoned him yesterday up

In Hungarian, on the other hand, verbal modifiers, including prefixes, must precede the finite verb in their neutral order. Therefore, I will assume that verbal modifiers in Hungarian move along with the finite verb, contrary to Dutch. Therefore, the Hungarian counterpart of (28b) is grammatical:

(29) János *fel*hívtaj őt tegnap t_j
John up-phoned him yesterday
'John phoned him up yesterday.'

The phenomenon that a verb takes along its prefix under movement is not so exceptional. Observe from (27) that this may also appear with V-raising. The following pair demonstrates this optionally applies with V-raising in Dutch as well:

(30) a. Ik heb [IP Jan *op* t_j] willen bellen$_j$ b. Ik heb [IP Jan t_j] willen *op*bellen$_j$
I have John up will phoned I have John will up-phoned

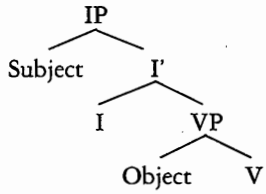
The reason for this dichotomy between V-movement and V-raising in Dutch is not clear to me. However, the Hungarian cases in which the prefix moves along can be accounted for much easier.

Suppose that the prefix may be incorporated by the verb at D-structure before movement applies. Incorporation of verbal modifiers is a quite general phenomenon in Hungarian (cf. the sections 4.4. and 5.3.1.). Hence, the tight connection beween the prefix and verb in the neutral order is a subcase of a much broader phenomenon. Furthermore, if incorporation takes place at D-structure, we expect that the complex verbs reflect the D-structure order. This appears to be the case. In all such cases, the verb is in final position.

Let us now turn to the consequences of the IP-parameter for the phrase structure of English, Dutch, Frisian, and Hungarian. Consider first English, a language with a strong I.

In languages with a strong I, an independent lexical I-item fills the I-position. This yields the following phrase structure for English:

(31) *English*



Observe from (31) that I governs the subject, and that V governs the object. Consequently, the minimal maximal domain of the subject does not coincide with the minimal maximal domain of the object. The domain of the subject is IP, the projection of its governor, whereas the domain of the object is CP, the projection of the verb.[15] Hence, we derive the following assumption:

(32) *Assumption 1:* In languages with *strong* I, the minimal maximal domain of the is IP but the minimal maximal domain of the object is CP

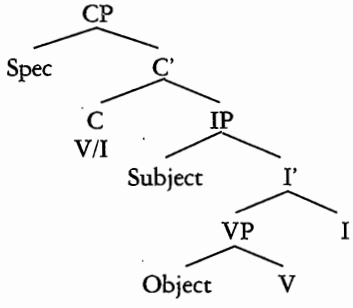
Let us now consider the phrase structure of languages with weak I. First of all, a weak I is a bound morpheme which has to satisfy the following condition:

(33) A bound morpheme may not remain stranded

This principle can be satisfied in several ways. Either I itself attaches to a host word, like C or V, or another lexical head, such as V, is moved to I. Let us examine how principle (33) is satisfied in Dutch, Hungarian and Frisian.

Bennis and Hoekstra (1987) have argued that in Dutch the V moves first to I before the V/I complex lands in C. Note that the merging of V and I in the I-position satisfies principle (33). A consequence of V-to-I movement is that I is lexically supported. Therefore, it may project into an IP (cf. (3)). Hence, Dutch has the following phrase structure:

(34) *Dutch*



(15) As regards the second claim, I will follow Koster (1987). According to Koster, CP and IP, are auxiliary projections to VP. This implies, among other things, that the local domain of categories governed by V is not necessarily VP but may be CP.

The V/I complex leaves a trace in I. By virtue of this trace, the verb governs the subject as well. Consequently, the minimal maximal domain of the subject is the same as the minimal maximal domain of the object, namely CP.

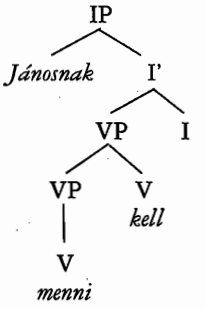
For Hungarian, I will assume that I-to-V movement satisfies principle (33). There are two pieces of evidence supporting this assumption. First, besides referential subject pro-drop, Hungarian also displays referential object *pro*-drop (cf. section 4.2.4.). According to Rizzi's (1986) theory on pro-drop, which I will follow here, referential overt pronouns may only be omitted if and only if they are governed by a Case-assigning head equipped with the relevant AGR-features. Hence, a proper context for object *pro*-drop can only be created if I lowers to V in Hungarian.

Second, infinitives in Hungarian may be optionally inflected for person-number agreement. However, this is only allowed in case the verbal governor does not host these features. For example, the auxiliary verb *kell* 'has to' may only be inflected for Tense but not for AGR (see, section 5.3.2. for details). Compare:

(35) Jánosnak menni(e) kell/kellett
John-DAT go-AGR3sg has to/had to
'John has/had to go.'

This sentence has the following D-structure:

(36)



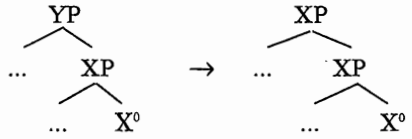
I cannot attach to *kell.* In order to avoid a violation of (33) two options are available. The infinitive moves to I, or I lowers to infinitive. The former option is ruled out, because of the HMC (cf. (6)). This principle forces *menni* to move through the position of the auxiliary but this is already lexically filled by *kell.* So I must lower to the infinitive to avoid a violation of (33). Hence, I-to-V movement derives the phenomenon of inflected infinitives in Hungarian.

Let us consider the implications of I-to-V movement for the phrase structure of Hungarian. I will assume that moved bound morphemes do not leave a trace. As a consequence, the IP-projection is bereft of its head. Following Chomsky (1973), I will assume that the minimal condition for domain *distinctedness* is that a domain must have a head. If this is not fulfilled, *L-containment* applies, which I will define as follows:

(37) *L-containment:* Projection XP *L-contains* projection YP if and only if YP directly dominates XP and YP does not contain a Y^0

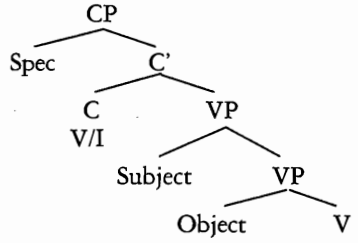
So, in the following configuration XP L-contains YP:

(38)



Hence, we derive the following phrase structure for Hungarian:

(39) *Hungarian*



Note from (39) that VP L-contains IP. As a consequence, the subject is adjoined to VP. The question then is what the governor of the subject is.

Following Chomsky (1982: fn.14), I will assume that an adjoined category is governed by the head of the category to which it is adjoined. This can be accommodated within Aoun and Sportiche's (1982) theory of *government* as follows:

(40) X *governs* Y if and only if
  a. X is an X^0 and
  b. X c-commands Y if X and Y are X", $X \neq Y$ and for V Φ a maximal projection, Φ dominates $X \rightarrow \Phi$ dominates Y where Φ includes all member-nodes of Φ

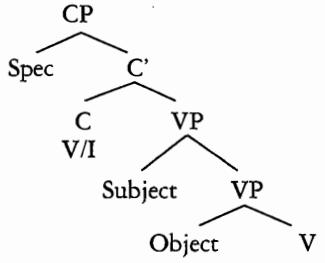
Therefore, in configuration (39) V governs the subject. This implies that the minimal maximal domain of the subject is the projection of V, that is, CP. Again, just as in Dutch, the minimal maximal domain of the subject is the same as the minimal maximal domain of the object in Hungarian.

Hoekstra and Marácz (1989) argue that I moves to C in Frisian independently of V. Note that this satisfies (33). Empirical evidence for I-to-C movement may be obtained from the fact that Frisian possesses inflected complementizers. Compare:

(41) Ik tink [$_{CP}$ *datst* [$_{IP}$ (do) jûn komst]]
  I think that-AGR2sg you tonight come-AGR2sg
  'I think that you will come tonight.'

If I moves independently to C without leaving a trace, then L-containment of the IP-projection applies, like in Hungarian. This yields the following structure:

(42) *Frisian*



Observe from (42) that in Frisian, like in Hungarian, the governor of the subject is V, and consequently the minimal maximal domain of the subject is CP.

In sum, there are several possibilities to satisfy principle (33) in languages with weak I. Either V moves to I, like in Dutch, or I moves to a host word, like C in Frisian or V in Hungarian. As a result, the subject in this type of language gets into the government domain of V. The minimal maximal domain of the subject is therefore identical with the mininal maximal domain of the object, namely CP. Hence, we derive the following assumption:

(43) *Assumption 2:* In languages with *weak* I, the minimal maximal domain of the subject is similar to the minimal maximal domain of the object, that is, CP

Both in languages with strong I and in languages with weak I the subject is structurally prominent over the object. However, in languages with weak I the minimal maximal domain of the subject is identical with the minimal maximal domain of the object, unlike in languages with strong I. In chapter five, I will argue that this covers the fact that subject-object asymmetries arise in both Dutch, English, Frisian, and Hungarian but that sometimes subject-object symmetries appear in Dutch, Frisian, and Hungarian where English displays subject-object asymmetries.

Let us consider now what the position of topicalized NPs, or quantifiers preceding Focus in the Hungarian phrase structure is (cf. 2.1.(28f)).

2.2.3. *CP is recursive within CP*

In the preceding section, I concluded that the Focus-position is identical to [Spec, CP]. If topicalized NPs and other quantifiers may precede Focus, then these phrases must be embedded under CP as well. Because of the fact that there may be infinitely many constituents in front of Focus, I will assume that CP is *recursive* within CP. This yields the following property of phrase structure in Hungarian:

(1) *CP is recursive within CP*

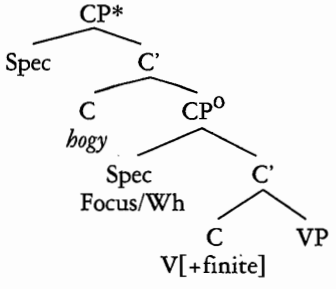
Let us consider whether we can find further empirical support for (1).

Indirect Wh-questions in Hungarian may be introduced by the complementizer *hogy* 'that'. For example, the verb *tud* 'know' may select a [+Wh] CP. Compare:

(2) Nem tudom hogy *kivel* találkozott János
Not know-AGR1sg that who-INSTR met-AGR3sg John
'I do not know who John met.'

Sentence (2) demonstrates that a complementizer may precede an indirect Wh-question in Hungarian, unlike in English. This demonstrates that the CP is recursive within CP in such embedded clauses:

(3)



Observe from (3) that the upper CP, CP*, is headed by the complementizer, and that the lower CP, CP°, serves as a landing-site for V-movement. Topicalized NPs can intervene between CP* and CP° requiring further recursions of CP:

(4) Nem tudom [CP* hogy [CP János [CP tegnap [CP° *kivel* találkozott]]]
not know-AGR1sg that John yesterday who-INSTR met-AGR3sg
'I do not know who John met yesterday.'

A consequence of (1) is that it also allows V-movement in embedded clauses with a lexical complementizer. Hence, we may say that V-movement in Hungarian is *generalized.* It does not only apply in root clauses, like in Dutch, but also in embedded clauses. Below I will demonstrate that V-movement in Frisian is sometimes also allowed in embedded clauses with a lexical complementizer. In that case, these clauses contain multiple CPs.

Property (1) of the Hungarian phrase structure is not so exotic as it looks at first sight. The phenomenon of multiple CPs has been attested in other languages as well. For example, it also appears in Spanish and Germanic.

Plann (1982) reports that in Spanish the complementizer *que* 'that' can occur before an indirect question after certain verbs of communication. In the following sentences, *que* precedes a Wh-phrase, similarly to Hungarian. Compare:

(5) a. Te preguntan [CP* que [CP° para qué quieres el préstamo]]
you ask-AGR3pl that for what want-AGR2sg the loan
'They ask you what do you want the loan for.'
b. Pensó [CP* que [CP° cuáles serían adecuados]]
thought-AGR3sg that which ones would be appropriate
'He wondered which ones would be appropppriate.'

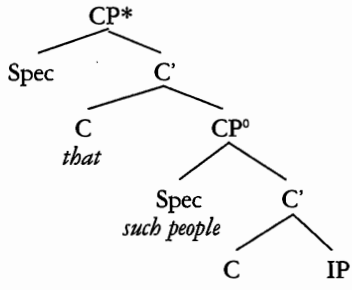
Let us discuss now some examples of multiple CPs in Germanic.

Hooper and Thompson (1973) have observed that the phenomenon of multiple CPs in English arises with embedded main clauses. Such clauses are embedded clauses to which root transformations in the sense of Emonds (1969) apply. For instance, objects may be topicalized in embedded main clauses:

(6) He said [CP* that [CP° *such people* [IP he doesn't like t]]]

This embedded clause may be represented in the following tree-structure:

(7)



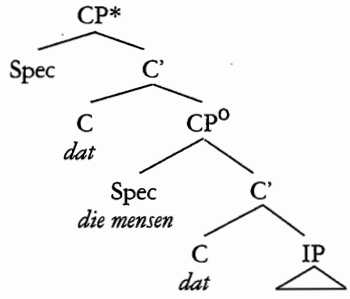
Note that [C, CP°] may not be filled in English by a lexical complementizer. The reason for this is that English does not allow a head of CP and its specifier to be fil-

led simultaneously (cf. Chomsky and Lasnik's (1977) doubly-filled COMP Filter). This is, however, a language-particular restriction, because in Dutch, for instance, topicalized objects may intervene between CP* and a CP^0 headed by a lexical complementizer. Compare:

(8) Ik denk [$_{CP*}$ dat *die mensen* [$_{CP^0}$ dat [IP die gek zijn *t*]]]
I think that those people that those crazy are
'I think that those people are crazy.'

This embedded clause has the following structure:

(9)



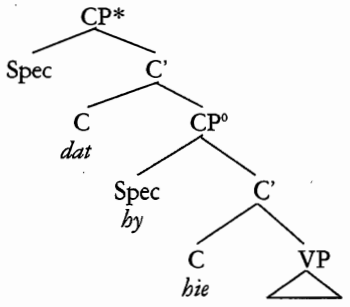
De Haan and Weerman (1986) discuss the occurrence of multiple CPs in Frisian. De Haan and Weerman note that V-movement is not only restricted to root clauses but may also apply in embedded clauses with a lexically filled C.[16] However, this is not allowed freely. Such embedded clauses must be complements of verbs denoting statements, feelings or observations.[17] Compare the following pairs:

(10) a. Ik leau [$_{CP}$ dat hy him wol rêdde kin]
I believe that he himself take care of can
'I believe that he can take care of himself.'
b. Ik leau [$_{CP*}$ dat [$_{CP^0}$ hy *kin* [him wol rêdde t]]]

(11) a. Ik sei [$_{CP}$ dat hy my sjoen hie]
I said that he me seen had
'I said that he had seen me.'
b. Ik sei [$_{CP*}$ dat [$_{CP^0}$ hy *hie* [my sjoen t]]]

Recall that the landing-site of a moved finite verb is C in Germanic. Hence, the (b)-sentences in (10) and (11) must contain a multiple CP. Therefore, the embedded clause of (11b), for instance, has the following structure:

(12)



(16) Hoekstra (1987) argues this is also the case in the Frisian Imperativus-pro-Infinitivo.

(17) De Haan (1983) reports that there are further restrictions on V-movement in clauses with a lexical complementizer. The matrix verb must be factual, it cannot be negative or modalized.

This phenomenon may also appear with adverbial degree clauses:

(13) a. Hy is sa meager [$_{CP}$ dat hy wol efter in reid skûlje *kin*]
he is so skinny that he behind a cane hide can
'He is so skinny that he can hide behind a cane.'
b. Hy is sa meager [$_{CP*}$ dat [$_{CP*}$ hy *kin* [wol efter in reid skûlje t]]]

De Haan and Weerman provide the following pieces of evidence which demonstrate that the clauses embedded under the complementizer of the (b)-sentences in (10), (11), and (13) have the same structure as root clauses.

(*i*) The distributional property of the subject clitic *er* 'he', a variant of the non-clitic *hy*. This clitic subject optionally appears after lexical complementizers such as *dat*, but it cannot appear sentence-initially. Compare:

(14) a. Hy sei [$_{CP}$ dat hy/*er* my sjoen hie]
he said that he me seen had
'He said that he had seen me.'
b. [$_{CP}$ Hy/**er* *hie* my sjoen t]
He had me seen
'He had seen me.'

Note that in clauses with a lexical complementizer and V-movement, the clitic subject *er* cannot appear immediately after the complementizer:

(15) Hy sei [$_{CP}$ dat [$_{CP}$ hy/**er* *hie* my sjoen *t*]]
He said that he had me seen

(*ii*) Consider the following sentence:

(16) Douwe [$_{CP}$ dy *woe* net komme *t*]
Douwe that wanted not come
'Douwe, he did not wanted to come.'

This sentence demonstrates that Left Dislocation may apply in root clauses. Note now that this phenomenon is blocked in regular embedded clauses (cf. (17a)) but it is allowed in embedded clauses with V-movement (cf. (17b)):

(17) a. *Hy sei [$_{CP}$ dat Douwe dy net komme woe]
he said that Douwe that not come wanted
b. Hy sei [$_{CP}$ dat Douwe [$_{CP}$ dy *woe* net kommen t]]
He said that Douwe that wanted not come

So embedded clauses with lexical complementizers and V-movement pattern the same as root clauses with Left Dislocation.

De Haan and Weerman conclude from these similarities that root clauses and embedded clauses with V-movement have, at least partly, a similar structure. According to De Haan and Weerman, this suggests that embedded clauses with a lexical complementizer and V-movement have a double CP. The upper CP is filled with the lexical complementizer, and the lowest CP serves as a landing-site for V-movement.

Hoekstra (1987) provides an additional argument for a double CP in such clauses.

(*iii*) Hoekstra observes that embedded clauses with a lexical complementizer and a moved verb have syntactic properties different from regular embedded clauses. The latter allow long Wh-movement of the object (cf. (18a)), whereas the former block this type of movement:

(18) a. *Hokker boek*$_i$ seist [$_{CP}$ dat se t_i lêzen hie]
which book said-you that she read had
'Which book did you say she had read.'
b. **Hokker boek*$_i$ seists [$_{CP}$ dat [$_{CP}$ se *hie*$_j$ [t_i lêzen t_j]]]

Hoekstra suggest that this difference is due to the fact that the embedded sentence in (18b) contains an extra CP-node which serves as a barrier in the sense of Chomsky (1986a).[18]

Recapitulating, in embedded clauses with lexical complementizers root transformations may apply involving Topicalization of objects (English, Dutch), and V-movement (Frisian). Furthermore, such clauses may contain indirect Wh-questions (Spanish). These phenomena require a recursion of CP. This demonstrates that multiple CPs appear in other languages than Hungarian as well, although it certainly is more restrictred in these languages.

It has been observed in the literature (cf. Hooper and Thompson 1973, De Haan 1983, and Plann 1982) that embedded main clauses exist only under certain conditions. First, these clauses must be introduced by the complementizer *that,* and second the verb of the matrix sentence governing the embedded main clause must have an asserted reading. The categorial component of Hungarian generates CPs much more freely. It must be admitted that the ultimate rationale behind this is unclear at the present state of research. I will leave this problem for further research.

A further consequence of (1) is that the universal condition on *scope-interpretation* (cf. Reinhart 1983) covers the descriptive generalization on the scope-interpretation of quantifiers (cf. 3.1.(28g)):

(19) A quantifier c-commands its scope at S-structure

The recursion of the CP within CP creates binary branching structures to the left of the verb which may accommodate the Quantifier Field. In a left-branching phrase structure, the leftmost constituent has the largest c-command domain. Therefore, in correspondence with (19), the leftmost quantifier in Hungarian has wide scope. A separate condition on scope in terms of linearity is thus superfluous (cf. also chapter six for discussion of scope phenomena).

2.2.4. Summary

I argued in this chapter that the underlying order of Hungarian is *SOV,* and that C is a strong and I is a weak governor in this language. Because strong governors must be lexicalized at S-structure, V-to-C movement applies. I presented empirical evidence for this type of movement involving the neutral SVO-order, V-second phenomena (Focussing, Inversion between finite verb and its modifier, and the phrase-structural position of Wh-phrases), and V-raising.

(18) In Chomsky's (1986b) theory, extraction out of embedded clauses with multiple CPs is allowed. The lower CP is not an argument. Therefore, a moved category may adjoin to it and subsequent movement of this category crosses one barrier only. Hence, no violation of Subjacency arises. The ungrammaticality of (18b) can, however, be accounted for under the assumption that the complementizer L-marks the lower CP. In that case, it becomes an argument and it can no longer act as an adjunction-site. Hence, extraction out of this category results in a violation of the Subjacency Condition. (See also chapter six for discussion of extraction with multiple CPs in Hungarian).

Further, I argued that CP is recursive within CP. Evidence for this was provided from various languages including Spanish, Dutch, English, Frisian and Hungarian. Repetition of CP appears unrestrictedly in Hungarian but not in the other languages. Multiple CPs accommodate indirect Wh-questions introduced by lexical complementizers, Topicalization and the position and interpretation of quantifiers. Furthermore, a recursive CP allows generalized V-movement. It may also apply in embedded clauses with multiple CPs. Hungarian shares this phenomenon with Frisian precisely when this language exhibits multiple CPs.

The phrase structure of Hungarian elaborated in this chapter resembles the phrase structure of Germanic languages, like Dutch, Frisian or German. C and I are the same in strength. C is strong and I is weak. We expect then that Hungarian and these languages will have several syntactic properties in common that are not shared by English. In this chapter, I concluded that languages with a strong C and a weak I display V-to-C movement. Furthermore, in chapter five I will show that some subject-object symmetries in Hungarian, including the absence of verb-object adjacency, the lack of that-trace effects, the absence of VP-deletion, and the lack of superiority effects, are caused by the fact that C is strong, and I is weak in this language. What is more, the very same properties turn also up in Dutch, Frisian, and German, but not in English.

3. THE LEXICON AND ASYMMETRIES

3.1. Introduction

This chapter discusses some properties of the lexicon and principles which mediate between lexical properties such as θ-assignment and syntactic structure.[1] I would like to argue for the following two claims:

I. In Hungarian the Unmarked θ-Assignment Conventions are operative

II. The realization of the Unmarked θ-Assignment Conventions is parametrized

Consider first the Unmarked θ-Assignment Conventions (cf. Carter 1967):

(1) *Unmarked θ-Assignment Conventions* (UTHAC)
   a. The theme role is assigned to the object GF
   b. The agent role is assigned to the subject GF

These conventions mediate between lexical properties of verbal predicators and syntactic structure. I will demonstrate that the assignment of θ-roles is guided by the principles in (1) in Hungarian. If these conventions are operative, then, this implies a *subject-object asymmetry,* that is, the subject and object GFs are discriminated structurally. This subject-predicate dichotomy of the sentence will be empirically supported by the following phenomena:

(*I*) An inventarization of the case frames which may be associated with basic verbal predicators in Hungarian. The cases selected by a verbal predicator and their corresponding θ-grids, that is, the set of θ-roles selected by that verbal predicator shows that Hungarian is a *nominative-accusative* language. The agent role of a basic verb is always related to the subject, i.e. the nominatively marked argument, and the theme role is always associated with the object, i.e. the accusatively marked argument. Hence, an interplay of the principles of Case theory, θ-theory, and the UT-

(1) See for studies of the lexicon in generative grammar: Bresnan (1982), Chomsky (1970; 1981), Guerssel et al. (1985), Hale (1983), Hale and Laughren (1983), Jackendoff (1972), Levin (1983), Marantz (1984), Ostler (1980), Perlmutter (1984), Simpson (1983), Stowell (1981), Williams (1981), and Zubizarreta (1985). These studies also discuss the relation between and the universal status of case-systems, θ-roles, and GFs.

HACs provide support for the subject-predicate dichotomy of the Hungarian sentence (cf. section 3.2.).

(*II*) Subjects may be assigned a θ-role *compositionally* but not objects. This is also the case in Hungarian, although ambiguities with predicates containing inalienable body part objects are absent from this language (cf. section 3.2.2. and 3.2.3. for discussion).

(*III*) Hungarian displays *transitivity alternations* such as the middle, unaccusative, ergative, and passive alternation (cf. section 3.3.). In spite of the fact that these alternations are *lexical* in nature, i.e. they can only be triggered by adding morphology to basic verbs, they provide evidence for a subject-predicate partitioning of the sentence. The following question then arises. Why are *syntactic* transitivity alternations in Hungarian absent but present in English?

Hale and Keyser (1985) argue that transitivity alternations are the result of the interaction of properties that enter into the lexical representation of basic verbs with both universal principles, such as formulated in Chomsky (1981), and language-specific rules. It may be clear that the absence versus presence of syntactic transitivity alternations with morphologically unaffected basic verbs in Hungarian and English respectively is due to a language-specific rule.[2] I will attribute this difference between Hungarian and English to a parameter, namely, to the θ-*Assignment Parameter:*

(2) θ-*Assignment Parameter* (THAP)
± apply the UTHACs in the syntactic representation of basic verbs

I will argue that Hungarian is specified positively for this parameter, whereas English may be specified negatively for it.

It has been claimed that the absence of syntactic transitivity alternations, i.e. NP-movement in Chomsky's (1981) sense, like syntactically derived middle verbs, ergatives, passives, and raising verbs is a diagnostic for a non-configurational sentence structure (cf. section 1.1.(6d) and É. Kiss 1987: 75). Since non-configurational languages do not distinguish the subject and object GF structurally, function-dependent operations cannot be executed in the syntax.[3] Below, I will demonstrate how the positive value of the THAP provides a straightforward answer to the question why in nominative-accusative languages, such as Hungarian, syntactically derived transitivity alternations might be missing in the syntactic representation of basic verbs. This will, then, compensate this diagnostic of non-configurationality *without* giving up the subject-predicate dichotomy of the sentence.

Let us, first, turn to a discussion of some properties of the lexicon. In section 3.2., I will introduce the subcomponents of which the lexical entries are composed.

3.2. The Structure of Lexical Entries

Following Hale and Keyser (1985), I will assume that in addition to its morphophonological and categorial features a lexical entry of a verb contains two parts

(2) See for transitivity alternations in other languages: Burzio (1981), Guerssel et al. (1985), Hoekstra (1984), Levin (1983), and Marantz (1984).

(3) Thoughout this study, I will adopt the position that GFs are structurally encoded (cf. Chomsky 1965: 68-74; 1981: 10).

which are relevant for its syntactic realization. The first part is the *Lexical Conceptual Structure* (LCS), roughly its dictionary meaning, from which the θ-*grid*, that is, the inventory of θ-*roles* can be derived (cf. Stowell (1981)). The second part is the subcategorization frame or *Lexical Structure* (LS), an abstract syntactic projection of the verbal lexical item, embodying the basic syntactic organization of its arguments. For example, the English dyadic verb *cut* has the following lexical entry:

(1) a. LCS for English 'cut':
{x produce linear separation in the material integrity of y, by sharp edge coming into contact with y}
b. θ-grid for English 'cut': (agent, theme)
c. LS for English 'cut': [$_s$ arg [$_{vp}$ arg v]]
(Hale and Keyser 1985: 16)

The entities in the θ-grid belong to a universal set of o-roles such as *agent, theme, goal, path,* etc. (cf. Gruber 1965, Fillmore 1968, and Jackendoff 1972). They are introduced by the participants involved in the action denoted by the verb. In the case of English *cut,* these participants are represented in the LCS of that verb by means of the variables *x* and *y.* These variables are projected into the θ-grid of the associated verb. In this way, for example, *x* and *y* of (1a) are represented, respectively, by the θ-roles *agent* and *theme* in the θ-grid (1b) of the verb *cut.*

In Chomsky (1981: 36) the following condition on the realization of θ-roles is formulated. Chomsky supposes that all θ-roles selected by a verbal predicator must be assigned to its arguments, the so-called θ-*criterion:*[4]

(2) θ-*Criterion* a. Each argument bears one and only one θ-role, and
b. Each θ-role is assigned to one and only one argument
(cf. Chomsky 1981: 36)

In (1c), the LS-projection of the transitive verb *cut* is depicted. Following, Hale and Keyser (1985), I will take (1c) to be a syntactic representation in the relevant sense. In particular, it represents the fact that the transitive verb governs an object, and that the subject is external to the VP.[5] With respect to its configurational properties I assume, in agreement with Chomsky (1981), that LS is universal.

The next question to answer is: how are the θ-roles in the θ-grid associated with the syntactic arguments in the LS of the verb? Hale and Keyser (1985) adopt the view that in syntactically nominative-accusative languages, verbs like transitive *cut,* which select both agent and theme θ-roles, assign the theme to the internal argument (the grammatical object) and assign the agent to the external argument (the grammatical subject). According to Hale and Keyser, this is the *unmarked* linking relation. They suppose that each of these conventions of θ-assignment is a genuine principle of UG, representing the unmarked case:[6]

(4) See for reformulations of the θ-Criterion: Hale and Laughren (1983), Higgingbotham (1985a), Rothstein (1983) and Williams (1983).

(5) The fact that the subject is always external to the VP is due to the operation of Predication. Compare Williams (1980) and Rothstein (1983) for details.

(6) Several authors, for example, Jackendoff (1972), Ostler (1980) and Carrier-Duncan (1985) have proposed a θ-hierarchy with a universal status:

(i) agent > theme > path (goal, source, location)

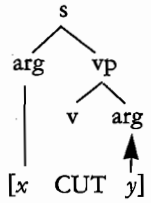
Observe that this hierarchy is rendered in structural terms by the UTHACs in (3).

(3) *Unmarked* θ-*Assignment Conventions* (UTHAC)
a. The theme role is assigned to the object GF
b. The agent role is assigned to the subject GF

These conventions become operative regularly where possible, i.e., where their application is not precluded for some reason, such as the prior application of some other convention or conventions, or the operation of other general principles of grammar.

The full lexical representation, which will be referred to as *Predicate Argument Structure* (PAS), of for example the verb *cut* indicates the projection of the agent and theme of the θ-grid, which are represented as the *x*- and *y*-variable respectively in LCS, onto the external and internal argument position of LS respectively. These connections will simply be indicated with the help of an association line. Compare:

(4) PAS for Transitive *cut*



The question is whether Hungarian is a nominative-accusative language. To answer this question, we will check whether the UTHACs are operative in Hungarian. This will be done by inventarizing the case frames and θ-grids associated with some basic verbs.

3.2.1. *Case Frames and* θ-*Grids in Hungarian*

Before an inventory of case frames and θ-grids in Hungarian can be presented, I will first discuss its case-system. Among Hungarian linguists, there is no agreement about which suffixes should be considered inflectional and which should be included into the set of case-markers.[7] The classification below follows Antal (1961b). According to Antal, case-markers are the markers that may be attached to each of the 14 different stems of the Hungarian noun. Consider:

(5)	*Case*	*marker*[8]	*Conjugation of the noun* fiú 'boy'
a.	NOM	-Ø	fiú (subj)
b.	ACC	-t	fiú*t* (obj)
c.	DAT	-nak/nek	fiú*nak* 'to the boy'
d.	INSTR	-val/vel	fiú*val* 'with the boy'
e.	ILL	-ba/be	fiú*ba* 'into the boy'
f.	SUBL	-ra/re	fiú*ra* 'onto the boy'
g.	ALL	-hoz/hez/höz	fiú*hoz* 'near the boy'
h.	INESS	-ban/ben	fiú*ban* 'in the boy'
i.	SUPER	-on/en/ön	fiú*n* 'on the boy'

(7) Kiefer (1988), for example, argues that Hungarian may have 18 cases if the following criterion is decisive:
(i) A bound morpheme is a case-marker if and only if it appears in a case frame.
(8) The alternants in this array are subject to the phonological rule of Vowel Harmony (cf. Vago 1980).

j.	ADESS	-nál/nél	fiú*nál* 'at the boy'
k.	ELAT	-ból/ből	fiú*ból* 'out of the boy'
l.	DELAT	-ról/ről	fiú*ról* 'of the boy'
m.	ABL	-tól/től	fiú*tól* 'from the boy'
n.	CAUS	-ért	fiú*ért* 'for the boy'
o.	TRANS/ESS	-vá/vé;-ul/ül	fiú*vá* 'become a boy'
p.	FORM/ESS	-ként;-képp;-ul/ül	fiú*ként* 'like a boy'
q.	TERM	-ig	fiú*ig* 'up to the boy'

Case-markers in Hungarian may have the following three main syntactic uses: (*i*) they may function as argument relators, (*ii*) as argument taking predicates (ATP), or (*iii*) as attribute relators.[9] In the use of argument relators, they mark the relation between an ATP and one of its arguments. The nominative and accusative cases are exclusively used as argument relators. The cases (c)-(q) may have both the function of argument relator signaling a thematically selected argument and of an ATP in which they subcategorize for a nominal complement yielding a 'free' or adverbial argument. The terminative case indicates that the NP to which it is attached is an adverbial argument.

I will discuss the use of case as attribute relator in section 4.6.[10]

Let us turn to an overview of the case frames and corresponding θ-grids of basic predicates in Hungarian appearing with respectively one argument (monadic), two arguments (dyadic), and three arguments (tryadic). I delay the introduction of derived predicates until section 3.3. Here I will not present a complete list of case frames but rather concentrate on the generalizations which may be derived from this sample.[11] Consider:

(6)

Verb	*Case frame* θ-*roles*
Monadic Verbs	
fut	'NOM run'
	agent
sétál	'NOM walk'
	agent
Dyadic Verbs	
lát	'NOM see ACC'
	agent - theme
szeret	'NOM love ACC'
	agent - theme
beszél	'NOM speak to DAT'
	agent - goal
vág	'NOM cut with INSTR'
	agent - instrument
megy	'NOM go into ILL'
	agent - goal

(9) See section 7.2.1. for the semantic-thematic classification of morphological case-markers in Hungarian.

(10) See Ackerman (1984) and Komlósy (1985) for further discussion of the functions of case-markers in Hungarian.

(11) A more extensive list of the Hungarian case frames can be found in: De Groot (1981a; 1984; to appear) Horvath (1983), Károly (1982), É. Kiss (1982a), Komlósy (1985), and Molnár (1966; 1967; 1973). These references discuss also some theoretical problems associated with case frames.

mutat	'NOM point to SUBL'
	agent - goal
áll	'NOM stand in INESS'
	agent (theme?) - goal
megy	'NOM go on SUPER'
	agent - path
áll	'NOM stand at ADESS'
	agent (theme?) - path
jön	'NOM come out of ELAT'
	agent - source
lelép	'NOM steps off from DELAT'
	agent - source
elfut	'NOM run away from ABL'
	agent - source
Tryadic Verbs	
ad	'NOM give ACC to DAT'
	agent - theme - beneficiary/goal
átköt	'NOM bind ACC with INSTR'
	agent - theme - instrument
bedob	'NOM throw ACC into ILL'
	agent - theme - goal
dob	'NOM throw ACC onto SUBL'
	agent - theme - goal
csatol	'NOM attach ACC to ALL'
	agent - theme - goal
akadályoz	'NOM hinder ACC in INESS'
	agent - theme - goal
kivesz	'NOM take ACC out of ELAT'
	agent - theme - source
beszed	'NOM withdraw ACC from ABL'
	agent - theme - source
beszél	'NOM speak to DAT about DELAT'
	agent - goal - source
beszél	'NOM speak with INSTR about DELAT'
	agent - instrument - source
beszél	'NOM speak to ALL about DELAT'
	agent - goal - source

If an agent is present in the θ-grid of a basic verb it is always associated with the nominative case, while a theme when present is always associated with the accusative case. These associations are rather fixed and they conform to a general rule of Hungarian grammar. For example, basic transitive verbs such as *lát* 'see' of the agent-theme semantic class occur always with a NOM-ACC case frame. This generalization is further supported by the association between case and θ-roles with active intransitive verbs. For example, the agent role of the active intransitive (unergative) verb *fut* 'run' is connected to the nominative case.

Languages in which the agent corresponds to the nominatively marked complement of a basic verb, while its theme to the accusatively marked complement are classified as nominative-accusative languages in the literature (cf. Marantz 1984:

198, among others). Hungarian, then, is a *nominative-accusative* language. These correspondences are established by an application of the UTHACs (cf. (3)). This implies that such languages recognize a structural subject-predicate partitioning since the GFs subject and object are defined as [NP, IP] and [NP, VP] respectively (cf. Chomsky 1965; 1986b). The structural configuration mediates between cases and θ-roles.

The nominative and accusative cases are assigned to the subject and object respectively by the Case-assignment rules for nominative-accusative languages (cf. Chomsky 1981; 1986b). Therefore, I will assume that Hungarian obeys the following Case-assignment rules as well:[12]

(7) *Case-Assignment Rules for Nominative-Accusative Languages*
a. Nominative Case is assigned to [NP, IP] under government by I[+AGR]
b. Accusative Case is assigned to [NP, VP] under government by V

As a result of the properties of nominative-accusative languages a classification of the overt morphological realization of Case in Hungarian may be set up. The nominative and accusative cases are structurally assigned under government and may henceforth be called *structural* Case, whereas the other cases in (5) are thematically dependent on verbal predicators.

(12) The nominative Case assignment rule (7a) is not general enough. It holds only from right to left. This follows from the fact that nominative Case may be assigned without a governing I[+AGR]. Koster (1986: 258) presents examples with nominative topics from German and Dutch. The topic positions in (ia) and (ib) are not governed by I[+AGR]:

(i) a. *Der Hans*, mit dem spreche ich nicht mehr
the Hans-NOM with him-DAT talk I not More
'Hans, I don't talk to him any longer.'
b. *Hij* een huis kopen, wie had dat kunnen denken
he-NOM a house buy who had that can think
'He buying a house, who could have imagined that.'

In Hungarian, too, nominative Case may appear without being governed by I[+AGR]. Consider, for example, the following two constructions:

(i) The complement of a nominalized verb is in the unmarked or nominative case (see also section 5.3.1.2. on Noun-Incorporation):

(ii) *fa* vágás
wood-NOM cut-NOMI
'wood-cutting'

(ii) Some PPs in Hungarian may display person-number inflection with pronominal complements (cf. section 7.3.1.). Consider:

(iii) *ő* mögöte
he-NOM behind-ppAGR3sg
'behind him'

The pronominal complement of these PPs bear nominativa case. This nominative Case assignment may be subsumed under (7a).

Nominal complements in such PPs appear also in the nominative case:

(iv) a fiú mögött
the boy-NOM behind
'behind the boy'

Note, however, from the minimal pairs in (iii) and (iv) that these complements do not trigger person-number inflection on the P.

Following Borer (1986), Taraldsen (1984), and Zwart (1988), I will assume that the nominative Case without being governed by I[+AGR] is a default Case. See chapter seven for the determination of the structural conditions on default Case in Hungarian.

For example, the three variants of the tryadic verb *beszél* 'speak' in (6) demonstrate that the case assigned to the non-nominative arguments is determined by thematic or lexical factors in a rather arbitrary way. Goal, instrument, goal and source correspond to the dative, instrumental, allative, and delative case respectively. A specific θ-role goes together with a particular case. Therefore, I will refer to the morphological cases in (5c)-(5q) as lexical case.

In order to formulate the principles of Case theory as strongly as possible it would be necessary to specify a unique structural position for lexical case as well. The determination of such a position and its relative structural prominence with respect to the positions of structural Cases is an empirical matter. In chapter five (see especially section 5.4.1.), I will return to these issues in more detail.

If the Hungarian sentence displays indeed a subject-predicate partitioning as is witnessed by the fact this language is a nominative-accusative language in which the UTHACs apply, it is to be expected that a subject-object asymmetry occurs with respect to the assignment of θ-roles. Chomsky (1981) suggests that *objects* (internal arguments) are assigned their θ-roles *directly* by their governing verb, whereas subjects (external arguments) are assigned a θ-role *compositionally* by the VP of which they are predicated. In the following section, I will discuss whether this asymmetry appears in Hungarian as well.

3.2.2. *The Asymmetric Nature of* θ*-Role Assignment*

Chomsky (1981: 104) has argued that a sentence like *John broke his arm* is ambiguous, depending on whether the subject bears the agent role or the patient role, in contrast to sentences such as *John broke the window* in which *John* has only an agent reading. Chomsky accounts for these readings by arguing that the subject but not the object may be assigned a θ-role *compositionally,* that is, by the VP.[13] The differences in the kinds of θ-assignments to the subject in the above sentences are clearly dependent on the choice of a different direct argument for the verb *break.*

Marantz (1984: 22-30) further elaborates on this asymmetry. Marantz presents two other pieces of empirical evidence for his hypothesis. First, he shows that simple transitive verbs in English express a wide range of predicates depending on the choice of the direct object but the predicates of transitive verbs remain unaffected by the choice of the subject. Second, Marantz argues that idiom frames in English are nearly always object-verb combinations but hardly ever of a subject-verb combination.

Evidence for a selectional subject-object asymmetry on the basis of the Hungarian equivalents of Chomsky's (1981) original examples cited above and the syntax of idiom frames does not easily carry over to Hungarian. Below I will attempt to make clear why compositional θ-assignment in Hungarian is more restricted than in English. However, discussion of idiom frames will have to wait until section 5.2.1.2.

Let us discuss first the *selectional asymmetry* between the (grammatical) subject of the predicate and the direct arguments of the verb.

(13) Jan Koster (personal communication) brings to my attention that facts about the world such as *his arm* and *the window* should not change rules of syntax. Although this position seems to me correct in essence, it must be noted that knowledge of the world such as 'agent of', 'theme of', etc. is mediated by θ-theory. Therefore, it should be not too surprising to find precisely in this domain interaction of structural conditions with knowledge of the world.

Horvath (1987) argues, convincingly in my view, that Hungarian exhibits selectional subject-object asymmetries. Horvath notes (cf. Horvath 1987: 150): "That selection of subjects by verb-object, but not selection of objects by verb-subject is quite systematically in Hungarian can be demonstrated by picking any common transitive verb, examining the variety of predicates it can produce with its objects, and contrasting this with the lack of parallel phenomena between the same verb and its subject". In order to support her claim, Horvath presents the following examples with lexical variants of the verb *vesz* 'take', i.e. *elvett* 'take away', *kivett* 'take out', and *átvett* 'take over' respectively. Compare:

(8) a. NP el*vette* a könyvet az asztalról
away-took the book-ACC the table-DELAT
'NP took the book from the table.'
b. NP el*vette* a pénzt
away-took the money-ACC
'NP accepted the money.'
c. NP el*vette* Marit
away-took Mary-ACC
'NP married Mary.'
d. NP el*vette* a kedvemet az utazástól
away-took the mood-npAGR1sg-ACC the trip-DELAT
'NP spoiled my interest in the trip.'

(9) a. NP ki*vette* a levelet a zsebéből
out-took the letter-ACC the pocket-npAGR3sg-ELAT
'NP took the letter out of his pocket.'
b. NP ki*vette* a szobát
out-took the room-ACC
'NP rented the room.'
c. NP ki*vette* a részét a munkából
out-took the share-npAGR-3sg-ACC the work-ELAT
'NP did his share of the work.'

(10) a. NP át*vette* a dijat
over-took the prize-ACC
'NP received the prize.'
b. NP át*vette* az irányitást
over-took the direction-ACC
'NP took charge.'
c. NP át*vette* a házifeladatot Marival
over-took the homework-ACC Mary-INSTR
'NP went through the homework with Mary.'
(Horvath 1987: 11)

Horvath notes that the θ-roles assigned to the subject NP in these sentences vary considerably, due to the wide range of predicates the verb produces with different objects (and other direct arguments). She proceeds to note that no corresponding variation in the interpretation of the object can be induced by varying the subject in the same case. The options for the assignment of θ-roles remain unaffected by the choice of subject:

(11) { A tanár / Egy bolond / Mindenki / Az apám / A csapat } { elvette / kivette / átvette } NP.

{ The teacher / A fool / Everyone / The father-my / The team } { away-took / out-took / over-took }

Obviously, Horvath presented clear instances of selectional subject-object asymmetries. From them it is apparent that the θ-role of the subject is affected by the choice of the direct argument of the verb but the choice of subject does not influence the assignment of the θ-role to the object of the verb. Therefore, it must be concluded that the subject-predicate partitioning of clauses is well-established and that Hungarian does not form an exception to this hypothesis.[14] Let us consider now compositional θ-assignment in English and Hungarian.

3.2.3. *Compositional θ-Assignment*

In the preceding section, I noted that some arguments concerning selectional subject-object asymmetries in English do not easily carry over to Hungarian. Consider again the sentences on which Chomsky (1981: 105) based an argument in favor of the idea that the VP assigns a θ-role to the subject of which it is predicated:

(12) a. John broke the window b. John broke his arm

According to Chomsky, the subject *John* in (12a) is normally understood as the active participant of the action. The sentence in (12b), however, has an additional interpretation, its more normal interpretation, in which *John* represents the passive participant in the sentence. Chomsky attributes this ambiguity to the fact that the subject is assigned a θ-role *compositionally* by the VP, i.e., by a combination of the verb and its direct arguments. The choice of the different internal arguments is responsible for the different readings in (12b). According to Hale and Keyser (1985), it seems to be quite generally the case in English that a VP of the form [V X's N], where *X* is an anaphor and *N* is a *body part,* can assign the *experiencer* role to the subject.[15]

Of course, the 'literal' reading of sentences like (12b) is also available. In this case the compositional θ-assignment of the VP to the subject is regular. The agent is associated with the subject subsumed under the UTHAC (3b).

From pairs as in (12), I conclude that the assignment of the agent role of a transitive verb to its subject might be *suppressed* in English in favor of the assignment of θ-role determined by the content of the predicate.

Let us consider the Hungarian equivalents of the sentences in (12):

(13) a. János eltörte az ablakot
John broke the window-ACC
'John broke the window.'

b. János eltörte a karját
John broke the arm-npAGR3sg-ACC
'John broke his arm.'

In contrast to the English pair, the subject *János* in the Hungarian sentences may only have an active reading.

In Hungarian, the two readings associated with the English (12b) are *disambiguated.* They are associated with two different lexical forms of the verb *eltör* 'break'. The active reading is expressed by the basic unaltered form *eltör* which is of the

(14) I will discuss some selectional subject-object symmetries in section 5.2.2. and I will evaluate their theoretical consequences in section 5.4.2.6.

(15) Chomsky (1981) refers to this θ-role as patient. Here I will follow Hale and Keyser (1985) in labelling this role as experiencer.

agent-theme semantic class and is associated with a NOM-ACC case frame. The passive reading associated with (12b) is expressed by employing the intransitive variant of the verb *eltör* by adding the verbal suffix *-ik,* as we will see below an instance of passive morphology, to the basic transitive verb stem *eltör.*[16] Compare:

(14) János karja eltörött
John arm-npAGR3sg broke

The verb in (14) has only *one* argument which is a possessive NP that consists of the experiencer, the possessor NP, and an inalienable body part, the noun-possessed. This possessive NP is marked nominatively.

The question is now: what should we conclude from the strategies employed by Hungarian in order to derive the readings of sentence (12b)? One could argue that the subject in Hungarian is not assigned its θ-role compositionally but relies on another kind of mechanism. This answer cannot be correct, however, because as we have noted in the preceding section the predicate of Hungarian transitive sentences may assign the subject a compositional θ-role as well. A more reasonable hypothesis is rather that a basic Hungarian transitive verb of the agent-theme semantic class realizes its θ-roles according to the UTHACs (cf. (3)). Therefore, the subject of a clause which contains a morphologically underived transitive verb of this semantic class receives always an active reading.

Obviously, the connection between the agent and subject is not so tight in English. It may be overruled by other grammatical factors. This dichotomy between English and Hungarian exemplifies that there is a *difference* in the application of the UTHACs between these languages. In the following section, I will discuss some instances in the domain of transitivity alternations which are due to this difference as well.

3.3. Transitivity Alternations in Hungarian

In the preceding section, I have presented evidence for the claim that the Unmarked θ-Assignment Conventions 3.2.(3), here repeated as (1),

(1) *Unmarked* θ-*Assignment Conventions* (UTHAC)
a. The theme role is assigned to the object GF
b. The agent role is assigned to the subject GF

apply in Hungarian. Recall that a language in which these principles hold is defined as a nominative-accusative language.

Suppose now that the UTHACs apply unrestrictedly in a particular nominative-accusative language. As a result of this, the D-structure thematic relations would be mirrored at surface structure. We expect then that in such a language no *transitivity alternations* would occur with morphologically unaffected basic verbs other than the ones made possible by the above rules. Transitive verbs of the agent-theme semantic

(16) The morpheme *-ik* itself appears only in the third person present tense: *eltörik* 'break-present tense-AGR3sg'. In the past tense, the transitive and the intransitive alternant can be kept apart, because they are conjugated differently. The transitive variant takes the definite conjugation, whereas the intransitive variant takes the indefinite conjugation. (See for a discussion of these verbal conjugations section 4.2.).

class will only have an agentive alternant, while non-agentive basic intransitive verbs cannot exist, because their subject position would remain empty yielding a violation of the universal requirement that all sentences must have a subject (the Extended Projection Principle of Chomsky 1982).

In this section, I will attempt to demonstrate that in Hungarian, as distinct from English, the UTHACs hold unrestrictedly. I will relate this difference to the following parameter, i.e. the θ-*Assignment Parameter:*

(2) θ-*Assignment Parameter* (THAP)
+/- apply the UTHACs in the syntactic representation of basic verbs

Hungarian takes the positive value of this parameter, whereas English may take its negative value. In Hungarian, the UTHACs apply whenever it is possible. In English, the application of these rules may be suppressed, although rule (1a) applies more rigidly than rule (1b). The theme role is nearly always associated with the object GF (but see section 3.3.6. on the Dative Shift Alternation), the realization of the agent role in English is more 'liberal'.

This parameter accounts for the fact why syntactic transitivity alternations, i.e. NP-movement in Chomsky's (1981) terminology, might be absent from the grammar of a *purely* nominative-accusative language. Because Hungarian is specified positively for (2), it is not possible to derive *syntactic* transitivity alternations which do appear in English, such as the Middle Alternation, the Causative/Inchoative Alternation, the Passive Alternation, Experiencer Verbs, Raising Predicates, and the Dative Shift Alternation. The difference in application of the UTHACs produces, then, superficial differences within the nominative-accusative languages yielding a typological difference, namely, the presence or absence of *NP-movement.*

Languages in which these conventions hold unrestrictedly, such as Hungarian, represent the *unmarked* case. The equivalents of the syntactic transitivity alternations in English can only be derived in Hungarian by carrying out a morphological operation which has the effect of *altering* the substructures in the lexical entry of a basic verb.

This section is organized as follows. First, I will discuss transitivity alternations which have a transitive and an intransitive alternant involving the *Middle Alternation* (cf. section 3.3.1.) and the *Causative/Inchoative Alternation* (cf. section 3.3.2.). It appears that in Hungarian the transitive alternant is always the basic one. Then, I will deal with transitivity alternations which can be derived in Hungarian only with the help of morphological markers, like the *Passive Alternation* cf. (section 3.3.3.), *Experiencer Verbs* (cf. section 3.3.4.) and *Raising Predicates* (cf. section 3.3.5.). Finally, section 3.4.3.6. will focus on the presence versus the absence of the *Dative Shift Alternation* in English and Hungarian respectively. This difference will be attributed to the fact that the theme role in English can be assigned by the structural position [NP, VP] but not in Hungarian.

3.3.1. *The Middle Alternation*

Consider the following sentences:

(3) a. John cuts the bread b. The bread cuts easily

Some basic transitive verbs like *cut, slice, kill, bribe, crush, assemble, maim, discourage, convince, corrupt,* etc. of the agent-theme semantic class may optionally undergo a process of *detransitivization* yielding the *Middle Alternation* (3b).[17] In the literature, two analyses are proposed for its derivation.

(I) Keyser and Roeper (1984) argue that this alternation may be derived from an interaction of Case and θ-theory. Some verbs are lexically specified to lose their ability to assign accusative Case to their object. In accordance with *Burzio's Generalization* which states:

(4) *Burzio's Generalization:* If some NP governed by V is assigned no Case, then the VP of which V is the head assigns no θ-role (cf. Burzio 1981)

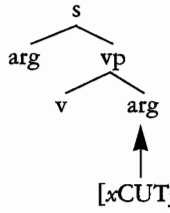
The verb or the VP fails to assign its subject a θ-role. Hence, for example, the D-structure object *bread* is moved to the subject position of (3b) in order to escape a violation of the *Case Filter.* The Case Filter is defined as:

(5) *Case Filter:* Every NP with phonological content must receive Case (cf. Chomsky 1983: 6)

(II) Hale and Keyser (1985) present an alternative analysis of the middle construction. Their approach is similar to that of Keyser and Roeper in that this alternation is the result of the interaction of independent modules and principles. According to Hale and Keyser, the external position in English is *not* a canonical agent position. The mapping of the agent onto the external position is optional, in contrast with the theme which is steadily linked to the D-structure object position. Hence, (1b) is not a core rule of English grammar.

They assign basic transitive verbs which participate in the middle alternation the following PAS (cf. also Guerssel et al. 1985):

(6) PAS for Middle *cut*



The theme role represented by the *y* variable in the LCS of the verb *cut* is associated with the internal position in LS by (1a). Hale and Keyser assume the projection of the agent role onto the subject position to be optional in such cases. In case the agent role represented by the variable *x* is assigned to the subject the Transitive Alternant (3a) is derived, whereas if it fails to project the Middle Alternant (3b) is derived. In the latter case, the theme argument is moved in the syntax to the subject position as an instance of the *Extended Projection Principle* which states:[18]

(17) The obligatory presence of an adverbial in the middle construction is not well understood yet. See Hale and Keyser (1985) for suggestions.

(18) This principle is responsible for the appearance of expletive *it* in the subject position of weather verb constructions (cf. (ia)) (see Chomsky 1981: 27, and in constructions with a preliminary subject *it* and a clause as a real subject (cf. (ib)) (see Stowell 1981):

(i) a. *It* rains b. *It* is clear *that he will come*

(7) *Extended Projection Principle* (EPP): Clauses must have subjects
(cf. Chomsky 1982: 10, Perlmutter 1984)

Under both analyses, the theme argument receives its Case-features in the subject position. This is, of course, only possible if the agent role is not present in the subject position, otherwise a violation of the θ-criterion would arise. Therefore, both analyses presuppose a relaxation of principle (1b) with respect to the realization of the agent role.

In section 3.2.3., I presented empirical evidence for the claim that the UTHACs apply unrestrictedly in the syntactic representation of basic Hungarian transitive verbs of the agent-theme semantic class. If that is correct, then we expect that there is no possibility in Hungarian for deriving syntactically middle constructions. This turns out to be the case. In order to derive this construction Hungarian necessarily employs an alternative strategy.

The transitive variant is always the basic alternant similar to English. *The Middle Alternation* is derived by *morphological* operations on these basic verbs. There are several morphological suffixes which have the effect of forming Middles. For example, the complex suffix *-ható* (cf. (8b)), which consists of a combination of the potentialis suffix (POT) *-hat* and the suffix of the participle present (pres.part.) *-ó,* or the reflexive suffix (refl) *-ódik* (cf. (9)) (see Károly (1982) for a classification of transitivity morphology in Hungarian):

(8) a. Janos vágja a kenyeret
John cuts the bread-ACC
'John cuts the bread.'

b. A kenyér könnyen vág*ható*
the bread easily cut-POT-pres.part.
the bread can be cut easily (lit.)
'The bread cuts easily.'

(9) a. Az emberek könnyen megvesztegetnek bürokratákat
the people easily bribe-AGR3pl burocrats-ACC
'People easily bribe burocrats.'

b. A bürokraták könnyen megvesztege*tödnek*
the burocrats easily bribe-refl-AGR3pl
'Burocrats bribe easily.'

An analysis for the *lexically* derived Middle Alternation in Hungarian may be elaborated along the lines of Chomsky (1981: 126). According to Chomsky (1981), morphological processes may absorb the assignment of a θ-role to the subject (for example passive morphology). Suppose, then, that the suffixes triggering the Middle Alternation have exactly this effect. They *absorb* the assignment of the agent role to the subject. Further, parallel to the English equivalents the theme argument in Hungarian is promoted to the subject position. This can be seen from the fact that it appears in nominative Case (cf. 3.2.(7a)). This movement to the subject position may then be the result of avoiding a violation of the Case Filter or the EPP.[19] Hence, the attachment of *passive* morphology to a basic transitive verb in Hungarian has the following consequences:

(19) Koster (1986; 1987: 262-266) argues that the obligatoriness of NP-movement in the case of passivization cannot be attributed to the Case Filter, because Case absorbed objects can remain in-situ in Dutch. According to Koster, this follows from the fact that the underlying object in passives appears both to the right and to the left of an immobile indirect object. In the former case it is in a VP-internal position. Consider:

(10) *The Properties of Passive Morphology*:
a. It absorbs the assignment of the agent role to the subject, and
b. The theme role is realized in surface subject position

Although Hungarian has no overt syntactic NP-movement, this rule may be triggered in the lexicon by adding passive morphology to a basic transitive verb.

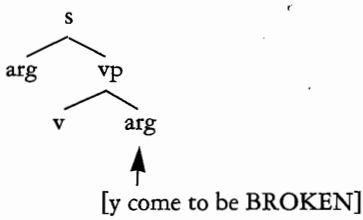
3.3.2. *The Causative/Inchoative Alternation*

Let us turn to the *causative/inchoative* alternation referred to in the theoretically oriented literature as 'ergative' alternation (cf. Burzio 1981) or 'unaccusative' alternation (cf. Perlmutter 1984). Some of the verbs belonging to this class are: *break, close, open, tighten, collapse, drop, slide, happen, arrive, appear,* etc. An example of the syntactic alternation at stake is provided by the following pair:

(11) a. The glass broke b. John broke the glass

The single argument in the intransitive alternant here denotes a passive participant in the event or process depicted by the verb (cf. Burzio 1981, Perlmutter 1984, among others). The theme role is assigned to the D-structure object in correspondence with (1a). Therefore, we may set up the following PAS of the verbs belonging to this class (cf. Hale and Keyser 1985 and Guerssel et al. 1985):

(12) PAS for Inchoative break



The NP bearing the object relation comes to bear the subject relation under the application of move-α. By the Case-marking rule 3.2.(7a) this argument is assigned nominative Case ensuring that the Case Filter is met. The theme argument is, of course, also the passive participant in the related transitive variant (11b).

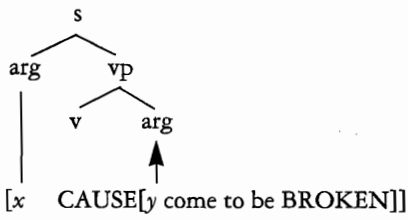
Jackendoff (1983) hypothesizes that the intransitive and transitive variant of this alternation are related by means of a *causative* rule. The principal observable effect of this rule in English is to embed the monadic LCS of the intransitive alternant as the complement of the causative function, which is itself dyadic, possessing an agentive argument as well as the complement it receives, as a result of the causativization process. Thus, for example, if the LCS of *break* is, roughly, [*y*BREAK], then the derived causative is, approximately, [*x*CAUSE(*y*BREAK)]. This rule is fully productive in English and applies to verbs which take an LCS of the form *y come to be a STATE*.

(i) a. dat_ *hem het boek* gegeven werd b. dat *het boek* hem__ gegeven werd
that him the book given was
'that he was given the book.'

Koster concludes that the obligatoriness of NP-movement in English cannot be caused by the Case Filter but by the EPP. This difference between English and Dutch, then, is related to the satisfaction of the EPP in these languages. In English NP-movement applies, whereas in Dutch the subject position may be filled optionally by the insertion of small *pro* (cf. Koster 1986).

Guerssel et al. (1985) assign the following PAS to the Causative Variant of the verb *break:*

(13) PAS for Causative *break*



Verbs participating in the middle alternation are basically *transitive* introducing an active and passive participant in the action denoted by them. The passive variant of the middle is derived by preventing the realization of the agent argument in the syntax. The inchoative alternation, on the other hand, is basically *intransitive* having a single passive participant.[20] The transitive variant of the inchoative construction, the causative alternant, is derived by adding an agentive role to the LCS of the inchoative verb which may then be projected onto the syntactic subject position.

In relation with the Hungarian equivalents of the examples in 3.2.(12), I already noted that the agent of a basically transitive verb in Hungarian may not be wiped out from its position by a rule assigning compositionally another θ-role to the subject. The verb involved in those examples belongs to the class of verbs which undergo the Causative/Inchoative Alternation. Recall that in Hungarian the Causative variant is the *basic* variant, that is, morphologically underived, whereas the Inchoative Alternant involves the suffixation of the passive morpheme *-ik*. Compare the Hungarian equivalents of the sentences in (11):[21]

(14) a. Az üveg eltörött
the glass broke-AGR3sg-indef
'The glass broke.'

b. János eltörte az üveget
John broke-AGR3sg-def the glass-ACC
'John broke the glass.'

In Hungarian as distinct from English, the 'Causative' Alternant is the basic alternant which realizes its agent and theme role according to the UTHACs, whereas the Inchoative Alternant is the derived one.[22] Its derivation involves the same effects as noted with respect to the derivation of the Middle Alternant (cf. (10)). The as-

(20) Compare Keyser and Roeper (1984) for the claim that there is an implicit agent in the middle construction but not in the inchoative (ergative) construction.

(21) See for the different types of conjugations in these sentence note 16.

(22) Hungarian has a morphological causative which may be derived by adding the suffix (CAUSAT) *-(t)at/(t)et* to verbal stems. If the verb is intransitive, then the original agent becomes the causee accusatively marked, while the causer is marked nominatively:

(i) a. János sétal
John walk-AGR3sg
'John walks.'

b. Mari sétál *tat*ja Jánost
Mary walk-CAUSAT-AGR3sg John-ACC
'Mary makes John walk.'

If the verb is transitive, then the original agent becomes the causee instrumentally marked, while the causer is marked nominatively in this construction as well (cf. (ii)):

(ii) a. A szabó varrja a ruhát
the tailor sew-AGR3sg the dress-ACC
'The tailor sews the dress.'

b. A szabóval varr*at*om a ruhát
the tailor-INSTR sew-CAUSAT-AGR1sg the dress-ACC
'I make the tailor sew the dress.'

See Hetzron (1976) for an extensive discussion of the morphological causative in Hungarian.

signment of the agent role to the subject is blocked and the theme role is realized in the surface subject position. Korponay (1980) and Károly (1982) observe that the following suffixes may yield Inchoative Alternations in Hungarian, like the reflexive suffix *-ódik/ődik, -ul/ül and -ad/ed.* These suffixes take a transitive base or a base unspecified for transitivity and add the syntactic properties in (10) to these stems.[23]

Compare an example with the suffix *-ódik.* The inchoative verb *becsukódik* 'close' (cf. (15a)) is formed by the suffixation of the morpheme *-ódik* to the morphologically unaffected variant *becsuk* 'close' (cf. (15b)):

(15) a. Az ajtó becsuk*ód*ott
the door closed-refl-AGR3sg
'The door closed.'

b. Mari becsukta az ajtót
Mary closed-AGR3sg the door-ACC
'Mary closed the door.'

In sum, the syntactic properties of the Hungarian equivalents of the Inchoative/Causative Alternation show that the transitive variant, unlike its equivalent in English of the agent-theme semantic class, is the *basic* alternant. The unmarked case involves the core case of the generalization in (1). The Inchoative Alternant is derived by morpholexical operations. Adding passive morphology results in the syntactic properties specified in (10).

A subgroup of the inchoative verbs (ergatives) is formed by the *unaccusative alternation.* An unaccusative verb, like an inchoative verb, assigns its θ-role to the object NP which appears as the surface subject under application of NP-movement. However, an unaccusative verb, unlike an inchoative, has no transitive counterpart.[24] Compare the following examples:

(16) a. Three men arrive b. A problem arises c. Several solutions exist

In Hungarian, Unaccusatives bear *passive morphology*, like the suffixes *-ik, -ódik/ődik, -ul/ül* and *-ad/ed.* Recall that passive morphology involves the syntactic properties in (10). (10a), which states that the agent role is not assigned to the subject, applies vacuously because these verbs are inherently monadic selecting only a theme role. Property (10b), however, also holds with Unaccusatives. This is clear from the fact that the underlying theme object appears as the surface subject, in the nominative Case (cf. 3.2.(7a)). Consider the Hungarian equivalents of the sentences in (16):[25]

(23) Hungarian has also a set of transitive morphology which has the oppositie effect of passive morphology. (i) The subject is assigned an agent role and (ii) a (verbal) root is turned into an accusative Case assigner. The suffixes of passive and transitive morphology often occur in oppositional pairs. For example *-ódik/ődik* (pass.) versus *-ít* (tr.), *-ul/ül* (pass.) versus *-ít* (tr.) and *-ad/ed* (pass.) versus *-aszt/eszt* (tr.). These suffixes may be added to transitive and intransitive verbal bases respectively, and to verbal bases which are unspecified for transitivity such as adjectives or nouns. Compare: *fehéredik* 'whiten' (pass.) versus *fehérit* 'make white' (tr.), *barnul* 'get brown' (pass.) versus *barnít* 'make brown' (tr.) and *fakad* 'spring' (pass.) versus *fakaszt* 'cause to spring' (tr.). See Bánhidi and Jókay (1960), Korponay (1980), Károly (1982) and De Groot (to appear) for more examples of such oppositions.

(24) See Perlmutter (1978) and Hale and Keyser (1985) for distributional differences between ergatives and unaccusatives. For example, Unaccusatives (cf. (ia)), unlike Ergatives (cf. (ib)) may participate in the rule of *There-insertion*:

(i) a. There arrived three guests b. *There closed a door

(25) See Szabolcsi (1986e; 1986f; and 1986g) for the participation of Unaccusatives in the definiteness effect in Hungarian.

(17) a. Három ember érk*ezik* b. Egy probléma ad*ódik*
three man arrive-AGR3sg a problem arise-AGR3sg
c. Több megoldás létez*ik*
several solution exist-AGR3sg

3.3.3. *The Passive Alternation*

So far I have investigated the effect of passive morphology on a transitive base or on an inherently intransitive base. It was concluded that the verbs derived in this way lack agentive arguments. The standard case of *Passivization* in English, however, allows the realization of an agentive argument as a *demoted* subject. Compare:

(18) a. The enemy beats the army b. The army is being beaten *by the enemy*

In (18a), we have the transitive verb *beat* of the agent-theme semantic class. By adding passive morphology *-en* to this verb the passive participle *beaten* is formed. The attachment of passive morphology has the properties in (10). The agent role is blocked from being assigned to the subject, and the underlying theme object is moved to the subject position where it can be assigned nominative Case by 3.2.(7a)). Note, however, that the agent role may be realized as a prepositional *by*-phrase.[26] The question arises whether passive morphology in Hungarian yields a two-place passive construction. In the literature, two instances of the passive alternation have been discussed.

(*I*) According to Károly (1982), the *Passive Alternation* is passing out of use but is entirely productive.[27] It may be formed by attaching to a transitive verb stem of the agent-theme semantic class the suffix (PASS) *-tatik/tetik*. Compare the Hungarian equivalents of the sentences in (18):

(19) a. Az ellenség megveri a sereget
the enemy beated-AGR3sg the army-ACC
b. A sereg megveretik az ellenségtől/ellenség által
the army beated-PASS-AGR3sg the enemy-ABL/ enemy by

The D-structure object is realized as the surface subject, and the demoted subject is expressed by a constituent marked ablatively or similarly as in English by a *by-phrase*.

(*II*) Kiss 1982a and Laczkó 1985a observe that *Deverbal Nominalization* with the suffix (NOMI) *-és/ás* patterns like passivization in Hungarian. Compare, for example:

(26) Compare Hoekstra (1986), among others, on the status of the demoted subject and passive morphology with passivization.

(27) Since the nineteenth century the passive formed with the suffix *-(t)atik/tetik* is not used any longer in active speech. To avoid this construction several kinds of strategies are employed. For example, it may be circumvented by either an impersonal infinitive construction:

(i) Itt dolgozni kell!
here work-INFI must
'Here one must work!'

or by a third person plural missing person construction (cf. section 4.2.4.1. for an analysis of this construction):

(ii) a. Ették a levest
ate-AGR3pl the soup-ACC
'The soup was eaten.'

b. Itt magyarul beszélnek
here Hungarian speak-AGR3pl
'Hungarian is spoken here.'

(20) a. Az ellenség megveri a sereget
the enemy beats-AGR3sg the army-ACC
'The enemy beats the army.'
b. A sereg megverése az ellenségtől/ az ellenség által
the army beat-NOMI-npAGR3sg the enemy-ABL/ the enemy by

Again, the D-structure theme is realized as the surface subject and the D-structure agent appears as a demoted subject assigned ablative case or as a PP.[28]

3.3.4. *Experiencer Verbs*

In section 3.3.2., I discussed the syntax of Unaccusative Verbs in Hungarian. These verbs lack an underlying agent subject. Two other verb classes in Hungarian also display this property, namely, *Experiencer Verbs* and *Raising Predicates.* The former are dyadic verbs which select a theme and an *experiencer* role and the latter are monadic verbs having a theme role in their θ-grids. Let us first consider Experiencer Verbs.

Compare, for example an experiencer verb construction with *tetszik* 'please':

(21) János tetszik Marinak
John please-AGR3sg Mary-DAT
'Mary likes John.'

Most of the Experiencer Verbs, like *tetszik* 'please', *hiányzik* 'is missing', and so on are *inherently* passive displaying passive morphology such as the suffix *-ik.* Consequently, they trigger (10b) as can be seen from the fact that the D-structure object theme is realized at surface structure in the nominative Case. The experiencer role is assigned to the dative complement (cf. Füredi 1976; É. Kiss 1982a and Pléh 1982).[29]

3.3.5. *Raising Predicates*

Raising Predicates in English as *seem, certain* and so on may select either an infinitival complement clause (cf. (22a)) or a tensed embedded complement (cf. (22c)):

(22) a. — seems [IP John to be sad] b. John seems [IP — to be sad]
c. It seems [CP that John is sad]

The embedded complement clause in (22a) lacks a fully specified I-node. According to Chomsky (1981), the D-structure subject in this clause cannot be assigned

(28) The suffix of the past participle *-(t)t* (cf. (ic)) follows the pattern of passive morphology, whereas the present participle *-ő* (cf. (ib)) leaves the transitive pattern of a verb of the agent-theme class unaffected (cf. (ia)) (see also Laczkó 1985):

(i) a. Az ellenség elszigetelte a hajót
the enemy isolated-AGR3sg the ship-ACC
'The enemy isolated the ship.'
b. A hajót elszigetelő ellenség
the ship-ACC isolate-pres part enemy
'The enemy isolating the ship.'
c. Az ellenségtől/ellenség által elszigetelt hajó
the enemy-ABL enemy by isolate-past part ship
'The ship isolated by the enemy.'
(Laczkó 1985: 93)

(29) Consider Pléh (1982) for the behavior of Experiencer Verbs with Switch Reference in Hungarian (cf. also section 5.3.4.).

Case, which results in a violation of the Case Filter.[30] However, the grammar provides for a way out of this conflict.

Raising predicates do not assign a θ-role to their subjects as can be seen from the realization of an expletive in (22c). Therefore the matrix subject position in (22a) qualifies as a landing site for NP-movement. So, the embedded D-structure subject of the infinitival complement may move to this position in order to receive a Case-feature of the matrix I. This results in the licit S-structure representation (22b).

As distinct from D-structure representation (22a), in (22c) a fully specified I is present in the embedded complement. Therefore, it assigns (nominative) Case to the embedded subject. An expletive is realized in the subject position of the matrix clause.

It has been observed that *Raising Predicates* in Hungarian, such as *látszik* 'seem' and *tünik* 'appear', do not trigger NP-raising (cf. É. Kiss 1987a: 64). Hence, the construction analogous to (22b) is lacking in Hungarian. In the equivalent of (22c), the Raising Predicate selects a tensed complement clause in which there is a null realization of the verb *van* 'be' (cf. Kiefer 1968) which acts as a nominative Case-assigner. Compare:

(23) (*Az) látsz*ik* [CP hogy János szomorú]
it seem-AGR3sg that John sad

From a closer investigation of some of the Raising Predicates it appears that morphologically they belong to the *-ik* class which we met already with Causative/Inchoative Alternation for example (cf. section 3.3.2.). As noted above, lexically raising predicates are monadic predicates which select a theme role. A result of the fact that they bear passive morphology is that this θ-role is realized in the subject position of these predicates. This can be seen from the fact that the anticipatory pronoun *az* 'that' which is a representant of the embedded clause in the LS of the matrix verb (cf. section 4.5.) appears in the nominative Case.[31] Therefore, a constituent cannot be raised from the embedded clause into the matrix clause without violating the θ-criterion (cf. 3.2.(2)).[32] Hence the absence of (22b) in Hungarian.

(30) Koster (1987: 262-265) argues that Raising Predicates are Unaccusatives in English and Dutch, and that the *it* expletive in raising predicate constructions (cf. (22c)) is an underlying theme object. This object represents the embedded clause in the LS of the verb. NP-movement of *it* to the subject position is then due to the EPP. In Dutch, the expletive *het* does not have to be present in raising predicate constructions:

(i) a. *Het* schijnt zeker dat Jan ziek is
it seems sure that John ill is
'It seems to be sure that John is ill.'

b. Dat Jan ziek is schijnt zeker

Similar to passivization, the subject position in Dutch may be filled optionally by smal *pro* (cf. note 19). These parallelisms between passive and raising predicates suggest that NP-movement is not caused by the Case Filter but rather by the EPP.

(31) Note, however, that in this case the anticipatory pronoun must be replaced by small *pro* (cf. chapter four).

(32) Raising Predicates may select not only an embedded clause but also a small clause:

(i) a. ____seems [s John sad]

b. John seems [s____ sad]

The difference with the infinitival complement in (22), is, that in these sentences the predicate is an AP instead of a VP. What (ia) shares with (22a) is that the embedded D-structure subject cannot be assigned Case because there is no suitable Case assigner present. By raising it to the matrix subject position it can receive a (nominative) Case from I. Compare the Hungarian equivalent of (ib):

3.3.6. *The Dative Shift Alternation*

The transitivity alternations discussed so far were all related to UTHAC (1b). In Hungarian, the agent is always connected to the subject, unless passive morphology specifies otherwise. In English, on the other hand, this mapping convention is not so strictly obeyed. Concerning the theme role, we have hypothesized so far that in both languages this θ-role is connected to the object. The question arises whether this is always the case. Marantz (1984) discusses the *Dative Shift Alternation* exemplified in the following pair:

(24) a. John gives a book to Mary b. John gives Mary a book

(24a) is an example of the unshifted alternant, whereas (24b) represents an instance of the shifted one.

Marantz presents the following analysis of this alternation. θ-roles may not only be assigned by lexical predicators and case-markers but also by *structural* positions. Further, Marantz assumes, adopting Chomsky's (1981) θ-criterion, that verbs may only assign one θ-role. He captures this restriction in his One role/One assigner principle. Marantz assigns the English verb *give* the following θ-grid:

(25) (*theme, goal*)

In (24a), the verb *give* assigns the *theme* role. According to the One role/One assigner principle, some other θ-role assigner must assign the *goal* role. Marantz argues that this is done by the preposition *to.* In (24b), however, the *goal* role is assigned by the verb. Hence, according to the One role/One assigner principle the *theme* role must be assigned by another θ-role assigner. Marantz claims (1984: 168) that the structural position [NP, VP] in English may qualify as a *theme* role assigner.

In Hungarian only the equivalent of the *unshifted* variant, that is, (24a) appears:

(26) János adja a könyvet Marinak
John gives the book-ACC Mary-DAT
'John gives the book to Mary.'

From the meaning of the verb *ad* 'give', it follows that this verb selects the same θ-grid as its English equivalent. *Ad* itself licenses the *theme* role, as is the case with the English variant (24a), whereas the dative marker has a similar function as the preposition *to,* namely, the assignment of the *goal* role.

The question is now of course: why is the shifted variant absent from Hungarian? If the analysis of the Dative Shift Alternation proposed in Marantz (1984) is correct, then, there might be two possibilities. Either Hungarian verbs do not assign a *goal,*

(ii) János szomorúnak látszik
John sad-DAT seem-AGR3sg

According to Komlósy (1985), the small clause complement is fully incorporated into the LS of the verb and forms a complex verb with it (cf. section 4.4.) in which the adjective is assigned dative case. It is unclear why Hungarian displays 'restructuring' in these cases (cf. also section 5.3.6.2. for the discussion of (ii) as an instance of secondary predication). Restructuring also applies with Raising Predicates which may select an infinitival complement (cf. Kálmán et al. 1984):

(iii) János futni látszott
John run-INFI seemed-AGR3sg
'John seemed to be running.'

or the structural object position [NP, VP] does not qualify as a licit *theme* role assigner. The first option is clearly incorrect as can be seen from the list in section 3.2.(6). Compare for example the verb *megy* 'go' which may select a goal argument:

(27) János a konyhába ment
John the kitchen-ILL went
'John went into the kitchen.'

Therefore, the latter option remains. The fact that the [NP, VP] position does not qualify as a θ-role assigner may be attributed to the strict application of the UTHACs. The [NP, VP] position in Hungarian may not be a *theme* role assigner because it is assigned this role *itself* whenever possible. Again, the application of such a convention seems to be more relaxed in English, although the association theme-object is more stable than agent-subject in that language.

3.4. Conclusions

In this chapter, I discussed some properties of the lexicon in general and the lexicon of Hungarian in particular. We have adopted the position that the lexicon contains several subcomponents such as LCS, LS, and a θ-grid. Further, we have adopted the Unmarked θ-Assignment Conventions, here repeated as (1), which mediate between lexical properties and syntactic structure:

(1) *Unmarked θ-Assignment Conventions* (UTHAC)
a. The theme role is assigned to the object GF
b. The agent role is assigned to the subject GF

From the assumptions of LS and these conventions, it follows that the sentence in Hungarian recognizes a *subject-predicate* divisioning.

This hypothesis has been supported, first, by the fact that Hungarian is a *nominative-accusative* language. Since the agent of underived transitive verbs of the agent-theme semantic class and of active intransitive verbs is associated with the subject, i.e. the nominatively marked argument, and the theme of underived transitive verbs of the agent-theme semantic class is associated with the object, i.e. the accusatively marked argument.

Secondly, the assignment of θ-roles is subject to a subject-object *asymmetry*. The subject but not the object may receive its θ-role *compositionally* in Hungarian as well.

Thirdly, Hungarian displays morpholexically induced *transitivity alternations*. These phenomena are instances of NP-movement which apply at D-structure. This can be seen from the fact that the D-structure theme object may appear in the subjective (nominative) Case of morpholexically derived Middles, Ergatives, Unaccusatives, Passives, Nominalizations, Experiencer Verbs and Raising Predicates.

I have further demonstrated that although Hungarian and English are both nominative-accusative languages, there are some *differences* in the domain of compositional θ-assignment by a predicate which contains an inalienable body part object, and in the domain of transitivity alternations.

In Hungarian, the subject may not be assigned the θ-role experiencer by a predicate which contains an inalienable body part with a transitive verb of the agent-

theme semantic class. The agent role of basic transitive verbs in Hungarian may not remain unrealized as in the English Middle Alternation, the agent role in Hungarian may not be introduced as with the case of the English Causative/Inchoative Alternation, and the theme in Hungarian may not be assigned by a structural [NP, VP] position as in the English Dative Shift Alternation.

These differences between Hungarian and English can be accounted for by a dichotomy in the application of the UTHACs. It is attractive to associate this dichotomy with a *parameter.* Intuitively, it is plausible to suppose that languages may display parametric variation in the way θ-roles and syntactic positions are related. Therefore, I will set the θ-Assignment Parameter as follows:

(2) θ-*Assignment Parameter* (THAP)
+/- apply the UTHACs in the syntactic representation of basic verbs

If we assume that Hungarian takes the positive value, of this parameter, and English may take its negative value the differences between these languages discussed above are accounted for. Thus, Hungarian is much stricter in the application of (1) in the syntax of basic verbs than English.

Transitivity alternations in Hungarian have in fact a fairly simple structure. The *core* cases are produced by the UTHACs, whereas the *alternants* such as Ergatives, Unaccusatives, Passives, Nominalizations, Experiencer Verbs and Raising Predicates are derived by adding passive morphology to the basic verbal stems. These morpholexical rules operate on the subcomponents of the lexical entries of these verbal stems, and have the effect of 3.3.(10).

It has been claimed that the absence of syntactic transitivity alternations (NP-movement), such as the lack of syntactically derived middle verbs, ergatives, passives, and raising verbs is a diagnostic for non-configurationality,[33] since in non-configurational languages the GFs subject and object cannot be distinguished structurally and hence function-dependent operations cannot apply in syntax. I have suggested, however, that a possible source for the absence of these alternations in nominative-accusative languages lies in the strict application of conventions (1).

Reineke Bok-Bennema (personal communication) points out to me that there is no one-to-one correspondence between overt syntactic NP-movement and the morphological encoding of transitivity alternations. According to her, in Spanish, for example, all transitivity alternations which are instances of NP-movement cooccur with a morphological reflex. Hence, it could be claimed that all morpholexically induced transitivity alternations in Hungarian are cases of NP-movement as well. Above I have shown that there is indeed some evidence for this hypothesis.

Rather, the problem of this chapter is formed by the following implication. If overt syntactic NP-movement is absent with transitive basic verbs, then it can only apply with the help of morpholexical means. This statement holds from left-to-right but not the reverse. Further, it also implies that a strict application of the UTHACs in a language *L* and the lack of the morphological means to manipulate them would

(33) The appearance of transitivity alternations in a particular language is an argument in favor of the configurational structure of such a language. Levin (1989) argues that Basque must have a subject-predicate dichotomy on the basis of the syntactic properties of Unaccusative Verbs in that language.

predict *L* to be active. Mary Laughren (personal communication) informs me that Warlpiri is such a case. It has no transitivity alternations such as Causatives, Passives, Anti-Passives, and so on. The only transitivity alternation appearing is the Causative/Inchoative Alternation, which is encoded morphologically.

4. THE PROJECTION PRINCIPLE IN HUNGARIAN

4.1. Introduction

It has been argued that the Projection Principle in non-configurational languages is satisfied only at LS.[1] In these languages, constituents may be base-generated freely at PS as a consequence of this parameter, and the relation between LS and PS may be either one-to-null, or one-to-many (cf. section 1.1.). The way in which the Projection Principle applies in non-configurational languages accounts for some of their properties, such as free word order, extensive use of null pronouns and split constituents.

In this chapter, I will present some empirical evidence for the hypothesis that the Projection Principle holds in Hungarian at *all* levels of representation. This implies that Hungarian is a configurational language and that the "non-configurationality" diagnostics above must be derived without making reference to a parametrization of the Projection Principle.

The intuitive sense of the Projection Principle may be stated as follows:[2]

(1) The θ-marking properties of each lexical item must be represented categorially at each level of representation: at LF, S-structure, an D-structure (cf. Chomsky 1982: 8)

In section 3.2., we pointed out that the θ-marking properties of each lexical predicator are associated with an LS. Therefore, this formulation of the Projection Principle may be replaced by (2):

(2) *Projection Principle:* The LS must be represented categorially at each level of representation (cf. Chomsky 1986a: 84)

(1) See Chomsky's (1981) parameter *Assume a GF*, Hale's (1983) *Configurationality* Parameter, Mohanan's (1983) distinction between *Lexical Structure* and *Configurational Structure* and Zubizarreta and Vergnaud's (1982) dichotomy between *Virtual Structure* and *Actual Structure* (cf. section 1.1. for discussion).

(2) See Chomsky (1981; 1986a), Bresnan (1982), Marantz (1984), and Pesetsky (1983) for discussion of the Projection Principle and its status within UG.

This principle specifies the relation between the PAS of a lexical predicator and its syntactic realization. The determination of this relation is a fundamental problem of any syntactic theory.[3] Note that (2) puts the strongest possible constraint on relations at different levels in the syntactic analysis of a sentence. The above formulation states that the relation between PAS and phrase structure is a *structure-preserving* isomorphism. Hence, syntactic configuration is projected from the lexicon. Consequently, the phrase structure rules become superfluous.

The relation between PAS and phrase structure has the following characteristics:

(3) a. identity b. biuniqueness c. obligatoriness d. locality

The Projection Principle determines that this relation is one of *identity*. The structural relations established by θ-assignment and subcategorization frames are preserved in the course of the derivation.

Identity between PAS and phrase structure does not affect word order. The linear ordening of constituents is relevant only at surface structure. Language particular directionality principles, like the *Head Parameter* (cf. Chomsky 1988) which specifies the order of heads and complements, yield surface word order.

The relation between PAS and phrase structure is *biunique* in the sense that each argument selected by a lexical predicator has precisely one counterpart in phrase structure. This excludes the possibility of having, for example, one-to-null or one-to-many relations. So, all the arguments of a lexical predicator are visible at surface structure.

Consider, for example, the following pair:

(4) a. John eats a cake b. John eats

Sentence (4a) contains the transitive verb *eat* of the agent-theme semantic class. In sentence (4b), the object NP is missing. The question arises now whether there is a null pronoun present in the phrase structure and whether the object NP is truly missing. The former option is ruled out by the fact that English is not a *pro*-drop language, it has no morphological means to license non-overt pronouns. The latter option is not allowed by the Projection Principle, since the mapping between PAS and phrase structure would be one-to-null in that case. From this it follows that verbs such as *eat* in English are specified in the lexicon as intransitive, and may optionally also be realized as transitive Vs.

The *obligatoriness* of the mapping between PAS and phrase structure has the following two consequences. Firstly, we observed that Hungarian has two types of cases (cf. section 3.2.1.), involving (*i*) structural Case (nominative governed by I[+AGR] and accusative governed by V) and (*ii*) lexical case, which is assigned under θ-government (cf. (3)-(16) of 3.2.(5)). Recall, furthermore, that both types of cases might function as argument relators indicating the dependency relation between the NPs which bear them and an argument taking predicates (ATP). As a consequence of the Projection Principle, NPs with these cases must be present at S-structure and surface

(3) Most linguistic frameworks incorporate something comparable to the Projection Principle. For example, in Montague grammar there is a homomorphism from syntax to semantics. This means that the mapping between semantic values and syntactic categories is structure-preserving (cf. Dowty et al.1981).

structure. Secondly, Chomsky (1985: 84) notes that if some element is "understood" in a particular position, then it is there in the phrase structure, either as an overt category that is phonetically realized or as an empty category assigned no phonetic form. This means that when NPs are missing from the phrase structure their position is filled by an empty category.

A further property of the relation between PAS and phrase structure is that it obeys a *locality* constraint. This constraint arises from the fact that the structural government relation between a head and its argument determines the LS. An NP in the phrase structure must be in the local domain of the verb of which the LS contains the argument to which that NP is related.

This locality requirement has consequences for the analysis of unbounded dependencies. For example, long Wh-movement fronts a Wh-phrase from its base-generated position in the embedded clause to the matrix sentence. The locality constraint on the mapping from PAS onto phrase structure dictates that in the embedded clause an empty category must be present which satisfies the θ- and subcategorization-features of the embedded verb. Empirical support for the local implementation of the Projection Principle will be postposed until chapter six, in which I will discuss long Wh-movement in Hungarian.

In this chapter, I will discuss the following phenomena from Hungarian bearing on the Projection Principle. These involve the *system of personal pronouns* (cf. section 4.2.), *Left Dislocation* (cf. section 4.3.), *complex verb constructions* (cf. section 4.4.), *embedded clause formation* (cf. section 4.5.) and *split constituents* (cf. section 4.6.). The properties in (3) characterizing the Projection Principle figure in all these phenomena.

Section 4.2. investigates the system of personal pronouns in Hungarian. This system is determined by a morpholexical and syntactic split between the nominative/accusative personal pronouns (pronouns assigned structural Case) on the one hand, and the personal pronouns with lexical case (cf. (3)-(16) of 3.2.(5)) on the other hand. The former have a constant lexical stem which is declined as an ordinary noun, whereas the latter have a stem which is often homophonous with the corresponding case-suffix. In order to derive a fully specified personal pronoun in these cases, person-number agreement must be added to the case-stem.

In accordance with the Projection Principle, the personal pronouns with lexical case may not be omitted when they function as a verbal complement. Personal pronouns assigned structural Case, however, are used for reasons of emphasis only and are preferably omitted in neutral contexts. Therefore, Hungarian is a so-called *pro*-drop language (cf. Chomsky 1981). As a consequence of the Projection Principle, an empty category must be present in the phrase structure of *pro*-drop languages. I will attempt to demonstrate that this empty category is small *pro* (cf. Chomsky 1982), because it displays the diagnostics of *pro:*

(5) a. It is recoverable from AGR
b. It is a non-anaphoric pronominal with independent (deictic) reference, and
c. It is free in its governing category

The Projection Principle is also operative in Left Dislocation. Section 4.3. shows that in Hungarian a pronominal item marks the complement position of a verb to which the left-dislocated NP is related.

Section 4.4. discusses complex verb constructions in Hungarian. The verbal prefixes involved are homophonous with personal pronouns bearing lexical case. The verbal prefixes receive an argumental interpretation if a verbal complement is selected.

Section 4.5. demonstrates that embedded sentences in Hungarian are always accompanied by a dummy pronoun which has a syntactic function comparable to expletive *it* in English. *It* holds the syntactic complement position of an embedded clause in order to satisfy the Case- and θ-features of a verb.

Section 4.6. analyzes *split* constituents in Hungarian. I will conclude that split constituents with NPs are highly restricted by syntactic and semantic conditions. If split constituents were not constrained, this phenomenon would constitute a counterexample against the Projection Principle. The mapping between PAS and phrase structure would be one-to-many in such cases.

4.2. The System of Personal Pronouns in Hungarian

In this section, I will discuss the system of personal pronouns in Hungarian and its relation to the Projection Principle. Not all the personal pronouns trigger the same conjugational pattern when accusatively specified. Hence, I will first have to introduce the two different conjugational patterns of the Hungarian verb, the so-called indefinite and definite conjugation.

4.2.1. *The Indefinite and Definite Conjugation of the Hungarian Verb*

Hungarian verbs may be conjugated with two different types of conjugations in all tenses and moods, the so-called *indefinite* and *definite* conjugation. Consider, for example, the indefinite and definite paradigm of the verb *lát* 'see' in the present tense:[4]

(1)

indefinite conjugation	*definite conjugation*
1sg. látok	1sg. látom
see-AGR1sg-indef	see-AGR1sg-def
2sg. látsz	2sg. látod
see-AGR2sg-indef	see-AGR2sg-def
3sg. lát ϕ	3sg. látja
see-AGR3sg-indef	see-AGR3sg-def
1pl. látunk	1pl. látjuk
see-AGR1pl-indef	see-AGR1pl-def
2pl. láttok	2pl. látjátok
see-AGR2pl-indef	see-AGR2pl-def
3pl. látnak	3pl. látnak
see-AGR3pl-indef	seeAGR3pl-def

The question arises: when are these patterns used? Roughly, the choice of these conjugational patterns depends on the *definiteness* feature of the *accusative* object of the verb. This may be captured by the following descriptive statement:[5]

(4) Hungarian personal suffixes are subject to Vowel Harmony (cf. Vago 1980).

(5) Szamosi (1976) argues that the indefinite pattern is the basic one and that the definite pattern is derived by Clitic Doubling. According to Szamosi, the definite pattern obeys the cross-linguistic condition on this rule. It takes place with all and only those direct objects which are definite.

(2) The definite paradigm is triggered in case the accusative object of the verb is definite, otherwise the indefinite paradigm is triggered

The next question to answer is: what counts as an indefinite or definite object? At this place I will not give an exhaustive answer to this question. The reason for this is that there is no unique criterion available to determine grammatical definiteness.

A classification of grammatical definiteness might be related to the (in)definite status of the entity denoted by the NP in the discourse. From this point of view two classes of objects may be distinguished. Firstly, objects which are indefinite or definite in the discourse, and consequently trigger indefinite or definite conjugation respectively. In these cases there is a perfect match between the (in)definite status of the entity denoted by the object and the conjugation it triggers. I will call NPs belonging to this group *properly (in)definite.* Secondly, the conjugational pattern triggered by an object cannot be related to the (in)definite status of the entity denoted by it. Such cases arise when it is impossible to determine whether a certain linguistic object denotes an (in)definite entity in the discourse, or when an NP connected to a definite entity in the discourse triggers indefinite conjugation and vice versa. Therefore, I will refer to the NPs in this class as *inherently (in)definite.* Consider first some examples of properly (in)definite NPs:

(3) *properly indefinite*
- NPs modified by the *indefinite* article *egy* 'a'
- *indefinite quantifiers: valaki* 'someone', *(egy) néhány* 'a few', and *semmi* 'nothing'
- *Wh-phrases: ki* 'who', and *mi* 'what'

properly definite
- NPs modified by the *definite* article *az* 'the'
- *quantifier: összes* 'all', *valamennyi* 'all of', *mindnyájuk* 'we', and *mindnyájatok* 'you all'
- *proper names: Mari* 'Mary', János *'John',* and so on
- *noun-possessed: az anyja* 'his mother', *az apja* 'his father', and so on
- *demonstrative pronouns: az ilyen* 'such', and *az a(z)* 'that'
- *reflexive pronoun: maga* 'himself'
- *reciprocae pronoun: egymás* 'each other'

Consider now some NPs which belong to the class of inherently (in)definite expressions:

(4) *inherently indefinite*
- *personal pronouns:* accusative 1sg, 2sg, 1pl, and 2pl
- *relative pronouns: aki* 'who', and *ami* 'which'
- *demonstrative pronouns: egy amolyant* 'one of that kind-ACC', *ugyanilyent* 'the same kind-ACC', and *egy ilyen* 'such'
- *quantifiers: mindent* 'everything-ACC'
- *demonstrative pronouns and universal quantifiers with partitive interpretation: azt* 'some of-ACC', *valamennyi* 'all' in the sense of 'some of'

inherently definite

- *personal pronouns:* accusative 3sg, and 3pl
- *relative and interrogative pronouns* ending in *-ik:*[6] *melyik* 'which', *valamelyik* 'someone', and *amelyik* 'whichever'
- *embedded clauses*

Compare the following pairs exemplifying the distribution of the indefinite and definite conjugation in Hungarian:

(5) a. Látok egy lányt
see-AGR1sg-indef a girl-ACC
'I see a girl.'

b. Látok valakit
see-AGR1sg-indef someone-ACC
'I see someone.'

c. Kit látok?
who-ACC see-AGR1sg-indef
'Who am I seeing?'

d. Látsz engem?
see-AGR2sg-indef me
'Do you see me?'

e. Látlak téged
see-AGR1sg-indef you-ACC
'I see you.'

a'. Látom a lányt
see-AGR1sg-def the girl-ACC
'I see the girl.'

b'. Látom Marit
see-AGR1sg-def Mary-ACC
'I see Mary.'

c'. Melyiket látom?
which-ACC see-AGR1sg-def
'Which one do I see?'

d'. Látom őt
see-AGR1sg-def him
'I see him.'

e'.Látom magamat
see-AGR1sg-def myself-ACC
'I see myself.'

Observe from the comparison between (5a) and (5a') that a properly indefinite and definite NP trigger the indefinite, and definite conjugation respectively. Moravcsik (1984) notes, however, that modification by a definite article is a *sufficient* condition for triggering the definite conjugation, whereas modification by the indefinite article is not always a sufficient condition for triggering the indefinite conjugation:

(6) a. Látom az egyiket
see-AGR1sg-def the one of-ACC
'I see one of them.'

b. Egy másikat is látom
an other of them-ACC also see-AGR1sg-def
'I see another of them also.'

The indefinite and definite conjugation is triggered also in the pair ((5b), (5b')) in which the properly indefinite quantifier *valaki* 'someone' and a properly definite name appear, respectively.

As noted above, however, in a number of cases there is no direct relation between definiteness in the discourse and the grammar. In the case of Wh-phrases there is even a *split* between *who*-phrases and *which*-phrases. Note from the comparison between (5c) and (5c') that accusative *who*-phrases trigger indefinite conjugation but accusative *which*-phrases trigger definite conjugation. According to Comrie (1975), the difference in the conjugation type between *who*-phrases and *which*-phrases in Hungarian is not controlled by definiteness in the strict sense, but by the related notion of

(6) Pronouns ending on *-ik* trigger the definite conjugation, except *másik* 'the other'. Compare:

(i) Kérek/*kérem másikat
ask-AGR1sg-indef/def other-ACC
'I want the other.'

restricted superset. In the case of *which*-phrases, the speaker presupposes that both speaker and hearer can identify the restricted set from which the choice is to be made, whereas with *who*-phrases this choice is completely free.

Pesetsky (1987) observes another split between *who*-phrases and *which*-phrases in English. *Which*-phrases in-situ fail to exhibit *superiority effects*, unlike *who*-phrases in-situ. Pesetsky relates this to the fact that *which*-phrases are *discourse-linked* whereas *who*-phrases are not. It would be worth exploring whether the split in the category of Wh-phrases in Hungarian is connected to discourse-linking.

In some cases the conjugational pattern triggered by the accusative object is the reverse of what we expect on the basis of relations in the discourse. The first and second person, i.e. the speaker and hearer in discourse, are referentially unique and hence count as *definite*. The third person, on the other hand, is assigned reference in discourse only. Therefore it counts as *indefinite*. Notice, however, that from a comparison between ((5d), (5e), and (5d'), it appears that exactly the opposite is the case concerning the conjugational-type. First and second person accusative objects trigger indefinite conjugation, whereas third person accusative objects trigger definite conjugation.

Summarizing, for our purposes it is sufficient to keep in mind that the conjugational pattern of the Hungarian verb is determined by the definiteness feature of the accusative object. By and large the descriptive statement in (2) captures the distribution of the indefinite and definite paradigm. Let us turn now to a discussion of the system of personal pronouns in Hungarian starting with the nominative and accusative personal pronouns.

4.2.2. *The Nominative/Accusative Personal Pronouns*

Consider the following paradigms:

(7) a. (Én) látom (őt) |*(őket)
I see-AGR1sg-def him/her|them
'I see him/her|them.'

b. (Te) látod (őt) |*(őket)
you-sg see-AGR2sg-def him/her|them
'You see him/her|them.'

c. (Ő) látja (őt) |*(őket)
he/she see-AGR3sg-def him/her|them
'He/she sees him/her|them.'

d. (Mi) látjuk (őt) |*(őket)
we see-AGR1pl-def him/her|them
'We see him/her|them.'

e. (Ti) látjátok (őt) |*(őket)
you-pl see-AGR2pl-def him/her|them
'You see him/her|them.'

f. (Ők) látják (őt) |*(őket)
they see-AGR3pl-def him/her|them
'They see him/her|them.'

(8) a. (Én) látlak (téged)/*(titeket)
I see-AGR1sg2sg/pl-indef you-sg/you-pl
'I see you.'

b. (Te) látsz (engem)/*(minket)
you-sg see-AGR2sg-indef me/us
'You see me/us.'

c. (Ő) lát-Ø (engem)/*(minket)|(téged)/*(titeket)
he/she see-AGR3sg-indef me/us |you-sg/you-pl
'He/she sees me/us/you.'

d. (Mi) látunk (téged)/*(titeket)
we see-AGR1pl-indef you-sg/you-pl
'We see you.'

e. (Ti) láttok (engem)/*(minket)
you-pl see-AGR2pl-indef me/us
'You see me/us.'
f. (Ők) látnak (engem)/*(minket)|(téged)/*(titeket)
they see-AGR3pl-indef me/us /you-sg/you-pl
'They see me/us/you.'

In (7) and (8), the full definite and indefinite paradigm of the transitive verb *lát* 'see' is listed. Overt pronouns marked nominatively and accusatively are used in Hungarian for reasons of emphasis only. In a neutral context, they are usually omitted. Mostly personal pronouns are recoverable from verbal inflection, which specifies person and number. Therefore, the behavior of these pronouns is subsumed by Chomsky's (1981) *Avoid Pronoun Principle*. Below I will return to an extensive discussion of the omissibility of overt nominative and accusative personal pronouns.

Let us first determine the *intrinsic* features of personal pronouns in Hungarian, that is the so-called ϕ-*features,* such as *number, gender*, etc. Observe from the glosses in (7) and (8) that personal pronouns are specified for *person* and *number,* similarly as their counterparts in English. Note, however, that personal pronouns in Hungarian are *not* specified for *gender.* The personal pronoun of the third person *ő* (cf. (7c), (8c)) may be translated in English with *he,* or *she.* This means they are neutral with respect to the feature gender. Further, recall that personal pronouns in Hungarian, in their accusative forms, are specified inherently for *definiteness* (cf. (4)). These pronouns of the first and second person are [+definite], and the personal pronouns of the third person are [+definite]. Hence, personal pronouns in Hungarian have the following ϕ-features:

(9) ϕ-*features of personal pronouns in Hungarian:* person, number, and definiteness

Let us turn to the personal pronouns corresponding to the cases (c)-(p) in 3.2.(5), i.e. the personal pronouns bearing lexical case.

4.2.3. *Personal Pronouns with Lexical Case*

Elsewhere (cf. Marácz 1984), I observed that the fully specified forms of the personal pronouns with lexical case differ from those of the nominative and accusative personal pronouns and other kinds of pronouns, like demonstrative pronouns, interrogative pronouns, etc. The latter group of pronouns have a constant lexical stem, i.e. the pronoun *itself,* which may be declined as an ordinary nominal such as *fiú* 'boy' in 3.2.(5). The stems of the personal pronouns with lexical case are, however, not constant lexical items but are often homophonous with the corresponding case-suffix:

(10) *Stems of personal pronouns with lexical case:*

DAT	*nek-*	SUPER	*rajt-*
INSTR	*vel-*	ADESS	*nál-*
ILL	*bele-*	ELAT	*belől-*
SUBL	*rá-*	DELAT	*ről-*
ALL	*hozz-*	ABL	*től-*

In order to receive a fully specified personal pronoun, person-number agreement must be added to the case-stems in (10). Compare, for example, the paradigms of the dative (cf. (11)) and sublative personal pronouns (cf. (12)):

(11) a. nekem
DAT-AGR1sg
'to me'
b. neked
DAT-AGR2sg
'to you (sg)'
c. neki
DAT-AGR3sg
'to him'
d. nekünk
DAT-AGR1pl
'to us'
e. nektek
DAT-AGR2pl
'to you (pl)'
f. nekik
DAT-AGR3pl
'to them'

(12) a. rám
SUBL-AGR1sg
'on me'
b. rád
SUBL-AGR2sg
'on you (sg)'
c. rá-Ø
SUBL-AGR3sg
'on him'
d. ránk
SUBL-AGR1pl
'on us'
e. rátok
SUBL-AGR2pl
'on you (pl)'
f. rájuk
SUBL-AGR3pl
'on them'

Obviously, case-stems are lexically specified for selecting AGR. Other lexical categories, such as nouns or the so-called dressed postpositions, also display this property (cf. chapter seven). AGR is "rich" enough to sanction the omission of an overt nominative pronoun in these constructions. An overt nominative personal pronoun is spelled out only when it expresses emphasis. Compare the paradigm of an inflected dative case-stem with the nominative pronouns:[7]

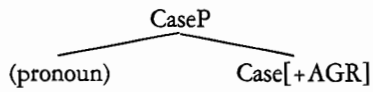
(13) a. (én) nekem
I-DAT-AGR1sg
'to ME'
b. (te) neked
you (sg)-DAT-AGR2sg
'to YOU (sg)'
c. (ő) neki
he-DAT-AGR3sg
'to HIM'
d. (mi) nekünk
we-DAT-AGR1pl
'to US'
e. (ti) nektek
you (pl)-DAT-AGR2pl
'to YOU (pl)'
f. (ő) nekik
they-DAT-AGR3pl
'to THEM'

Therefore, nominative personal pronouns in combination with inflected case-stems behave like nominative and accusative personal pronouns in combination with verbal agreement. In neutral contexts, they are omitted, and they are recoverable from person-number agreement. Therefore, I will assume that the nominative personal complements of a case-stem fall under the Avoid Pronoun Principle as well. (See the following section and chapter seven for a discussion of the syntactic role of AGR

(7) The nominative third person plural pronoun (cf. (13f) is homophonous with the nominative third person singular pronoun (cf. (13c)). Regularly, the nominative third person plural pronoun is *ők* 'they'. The omittance of *-k* in (13f) is due to a functional principle of redundancy, because AGR already marks plurality. This phenomenon appears also in inflected NPs and PPs (cf. section 7.3.1.).

in Hungarian). The projection of an inflected case-stem with a nominative personal pronoun may be represented in the following tree-diagram:

(14)



Anna Szabolcsi (personal communication) points out that the existence of the forms in (13) make possible an alternative analysis of personal pronouns with lexical case. Instead of taking the case-suffix as the stem of a personal pronoun with lexical case, it would be a *regular* case-ending attaching to the constant lexical stem provided by the personal pronoun marked nominatively. Note that under this proposal, personal pronouns with lexical case would have a morphological structure identical to other inflected lexical items such as nouns. There are, however, two arguments against this position.

First, if a nominative personal pronoun is the stem of the lexical case forms of personal pronouns instead of a case ending, it would be unclear why the suppletive forms *rajt-* and *belol-* of the superessive and the elative case appear with fully inflected forms and not the corresponding regular case endings *-n* and *-ből*. So, why do we not find (15b) and (16b) instead of (15a) and (16a)?:

(15) a. (én) rajtam
I SUPER-AGR1sg
'on me'
b. *énen
I-SUPER

(16) a. (én) belőlem
I ELAT-AGR1sg
'out of me'
b. *énből
I-ELAT

Second, Vago (1980: 97) formulates a phonological rule capturing the behavior of the *v-* of the instrumental case-suffix *-val/vel* and the translative suffix *-vá/vé*. The initial sound of these suffixes undergoes *total assimilation* with a preceding consonant of the stem to which these suffixes are added:

(17) *v-Assimilation*
C + v
1 2 3 → 1 2 1

Compare, the phonetic realization of the instrumental form of the noun *vén* 'old one':

(18) vén + *-vel* → vén*nel*

If the nominative personal pronoun *én* 'I' were the stem of the instrumental personal pronoun, the conditions of *v*-Assimilation would be met. We would expect, then, the initial *v-* to assimilate to the preceding *-n,* resulting in:

(19) én + -vel → *énnel

However, this form does not exist. The grammatical form *énvelem* (I INSTR-AGR1sg) 'with me' suggest that not the personal pronoun but the case-marker is the stem.

If we compare the case-markers in 3.2.(5) with the personal pronouns with a case-stem in (10), it appears that the *translative, essive, formalis,* and *terminative* are ab-

sent in the latter. Thus, we do not find the following personal pronouns, among others:

(20) a. *(én) kéntem
I TRANS-AGR1sg
'becoming me'

b. *(én) ülöm
I ESS-AGR1sg
'like me'

c. *(én) vém
I FORM-AGR1sg
'like me'

d. *(én) igem
I TERM-AGR1sg
'until me'

According to Komlósy (1985), the primary function of these cases is to mark *secondary predication.* They indicate that the arguments to which they are attached are referentially bound to another argument of the predicate. NPs with translative, formalis, or essive function as constants with an idiomatic sense (cf. section 5.3.6.2.). An intrinsic property of personal pronouns, however, is that they may have independent reference. Hence, this explains why these cases do not have pronominal forms. The terminative case is the only case-suffix which marks exclusively *non-selected,* adverbial NPs. The cases which serve as stems for personal pronouns, however, may function both as argumental and adverbial case. Obviously, this is a necessary condition for being a member of the group in (10). This accounts, then, for the fact that the personal pronouns of the terminative case do not exist.

Paradigms (7) and (8) in the preceding section demonstrate that the personal pronouns of the nominative and accusative may be omitted. The question arises whether this occurs with the personal pronouns with lexical case as well. Consider the following examples with argumental lexical pronouns:

(21) a. Beszélek *(neki)/*(nekik)
speak-AGR1sg he-DAT/they-DAT
'I am speaking to him/them.'

b. Várok *(rá)/*(rájuk)
wait-AGR1sg he-SUBL/they-SUBL
'I am waiting for him/them.'

c. Találkoztam *(vele)/*(velük)
met-AGR1sg he-INSTR/they-INSTR
'I met him/them.'

The verbs *beszél* 'speak to', *vár* 'wait for', and *találkoz* 'meet' may select an inherent dative, sublative, and instrumental case, respectively. Observe that the pronominal forms of the lexical cases may *not* be dropped. This contrasts, as we have noticed above, with the behavior of nominative and accusative personal pronouns.

The personal pronouns in the sentences (21a) and (21b) may sometimes be omitted. However, in those cases the meaning is not preserved. So, if the personal pronouns are dropped these sentences mean *I am speaking* and *I am waiting.* The reason that verbs such as *speak* or *wait* may have two grammatical variants is related to the fact that these verbs may be specified in the lexicon both as transitive and intransitive. Therefore, in case the pronominal forms of lexical cases are missing, they are truly missing. Let us turn now to a discussion of the conditions on the omission of personal pronouns in Hungarian.

4.2.4. *Pro-drop in Hungarian*

In this section, I will discuss the restrictions on omission of personal pronouns, that is *pro-drop,* in Hungarian. The question arises whether the syntactic position of a

dropped pronoun remains empty or is filled by a *null* pronominal. Chomsky (1982) identifies the missing pronominal in such cases as the empty category small *pro*. According to Chomsky, *pro* has the following properties:

(22) a. It is recoverable from AGR
b. It is a non-anaphoric pronominal with independent (deictic) reference, and
c. It is free in its governing category

The presence of null pronominals is guaranteed in case of *pro*-drop by the Projection Principle in combination with the θ-criterion. These principles are supported empirically if evidence can be provided for the claim that the non-overt counterpart of a full pronoun is present in the syntactic representation. First, I will attempt to demonstrate that in case personal pronouns are dropped in Hungarian, *pro* is actually present (cf. section 4.2.4.1.). After we have settled this, I will formulate the conditions on the distribution of *pro* in Hungarian (cf. section 4.2.4.2.).

4.2.4.1. *Is There pro in Hungarian?*

Above we noted that nominative and accusative pronouns are usually omitted in a neutral context (cf. the paradigms (7) and (8)). Consider again clause (7a), here repeated as (23):

(23) (Én) látom (őt)
I see-AGR1sg-def him/her
'I see him/her.'

First of all, observe that omission of the overt pronoun does not affect the interpretation of the clause. This implies that a non-overt item with independent deictic reference must be present in the syntactic position of the overt pronoun.

Of course, one could argue that verbal agreement takes over this function of personal pronouns when they are omitted. Hence, I will present more sophisticated evidence for the presence of a null pronoun in the case of *pro*-drop. This evidence comes from: (*I*) the *parallel* distribution of overt and null pronominals (with syntactic principles such as the Binding Principles), and (*II*) the *different* distribution between overt and null pronominals in various syntactic phenomena. Let us first turn to a discussion of the cases in (*I*).

(*I*) Recall that binding theory specifies the relation of referential expressions to possible antecedents. The conditions on which I will rely in the argumentation below are the *Binding Principles B* and *C* (cf. Chomsky 1981: 188). These conditions specify the environment in which a pronominal and a name may be bound:

(24) a. *Binding Principle B*
A pronominal (a category that may be referentially independent or may depend upon an antecedent for its reference, and thus includes the classes of pronouns) is free in its governing category
b. *Binding Principle C*
An R-expression (a category that is referentially independent, and it includes all other NP-types, for example names and Wh-traces) is free

A parallel distribution between an overt pronoun and *pro* shows up with structural conditions on *coreferentiality* between (*i*) a *pronoun* and another *pronoun* or *name,* and (*ii*) between a *pronoun* and a *Wh-trace.* Let us first consider (*i*).

(*i*) Compare the following sentences:

(25) a. *(Ő) látta (őt)
he saw-AGR3sg him
*'*He/she* saw *him/her.*'

b. *(Ő) látta *Marit*
she saw-AGR3sg Mary-ACC
*'*She* saw *Mary.*'

c. **Mari* látta (őt)
Mary saw-AGR3sg her
*'*Mary* saw *her.*'

d. (Ő) látta az (ő) anyját
she saw-AGR3sg the she mother-npAGR3sg -ACC
'She saw her mother.'

e. Az (ő) anyja látta (őt)
the she mother-npAGR3sg saw-AGR3sg her
'Her mother saw *her.*'

The ungrammaticality of a coreferential reading in the clauses (25) is accounted for by either Binding Principle B or C. *Disjoint reference* in the English equivalents of the clauses (25a)-(25c) is covered by Binding Principle B (cf. (25a), and (25c)) and Binding Principle C (cf. (25b)). In (25a) and (25c), the object pronoun is bound in its governing category, that is the sentence, and in (25b) the name in object position is not free, because it is bound.

The Hungarian counterparts exemplifying disjoint reference may be ruled out with the help of the Binding Principles in a *similar* fashion. Observe now that with respect to the coreferential interpretations in (25a)-(25c) there is no substantial difference between an overt and non-overt pronoun. This suggests that if overt pronouns are dropped in Hungarian null pronouns are present at their positions in syntax.

The pairs in (25d)-(25e) illustrate a similar point. Both the subject and object pronominals and the pronominals embedded in the possessive NPs in (25d) and (25e) are *free* in their governing categories. The clause counts as the governing category for the subject and object pronominals, and the possessive NP counts as the governing category for the embedded pronominals (cf. section 7.4.2.3. for this claim). So, a grammatical reading under coreferentiality of the personal pronouns is allowed by Binding Principle B both in (25d) and (25e). The grammaticality of these sentences remains *unaffected* in case one of the overt pronouns or both overt pronouns are omitted.

(*ii*) Horvath (1987: 140) presents an argument for the presence of pro and Wh-*trace* in the syntactic representation based on Binding Principle C. Horvath discusses the following pair:

(26) a. **Kitől* gondoltad hogy (ő) gyanította hogy Mari
who-DELAT think-AGR2sg that s(he) suspected-AGR3sg that Mary
ellopott egy könyvet t?
stole-AGR3sg a book-ACC
*'From *who* did you think that *s(he)* suspected that Mary had stolen a book t?'

b. *Ki* mondta hogy (ő) gyanította hogy Mari ellopott
who said-AGR3sg that s(he) suspected that Mary stole-AGR3sg
tőle egy könyvet?
he-DELAT a book-ACC
'Who said that *s(he)* suspected that Mary had stolen a book from *her/him?*'

Horvath argues that (26a), unlike (26b), is a case of *Strong Crossover* (SCO), that is, a Binding Principle C violation with Wh-traces. According to Horvath, the trace of

Wh-movement in the deepest embedded clause of (26a) may not be coindexed with a pronoun in the intermediate clause. This follows from the requirement that Wh-traces fall under Binding Principle C, and thus have to be free. The grammaticality of (26a) and (26b) does not change in case the pronouns in the intermediate clauses are dropped. Horvath concludes, then, that SCO yields an argument for both Wh-trace and small *pro* in Hungarian (cf. the sections 5.2.3. and 5.4. for SCO effects).

(*II*) The null-hypothesis is that non-overt pronouns have the same set of ϕ-features as overt pronouns. However, I will demonstrate that overt pronouns in Hungarian clearly have *different* grammatical features than non-overt pronouns. If correct, then, this would provide an argument for their independent existence in the grammar. I will illustrate this by investigating the following phenomena. including (*i*) the *weather verb construction,* (*ii*) coreference with *third person pronouns,* and (*iii*) the *impersonal passive construction.*

(*i*) Consider an instance of a *weather verb construction* in English:

(27) *It* rains

The Extended Projection Principle (cf. 3.3.(7)) is responsible for the appearance of expletive *it* in this type of construction (cf. Chomsky 1981: 27). With Chomsky (1981: 325), I will assume that weather verbs assign a quasi-ϕ-role to their subject NP. In Hungarian, overt expletives such as the demonstrative pronoun *az* 'that' may not appear in weather verb constructions.[8] Compare:

(28) a. (**Az*) esik
that rains
'It rains.'

b. (**Az*) zuhog
that pours
'It is pouring with rain.'

c. (**Az*) locsog
that plashes
'It is plashing with rain.'

d. (**Az*) villámlik
that lightens
'It is lightning.'

e. (**Az*) dörög
that thunders
'It is thundering.'

The question arises now whether there is a non-overt expletive pronoun present in the syntactic representation of these phrases. The noun undergoing the action in weather verb constructions may be spelled out in Hungarian (cf. Molnár 1967):

(29) a. Esik az *eső*
rains the rain
'It rains.'

b. Zuhog az *eső*
pours the rain
'It is pouring with rain.'

c. Locsog az *eső*
plashes the rain
'It is plashing with rain.'

d. Villámlik az *ég*
lightens the sky
'It is lightning.'

e. Dörög az *ég*
thunders the sky
'It is thundering.'

(8) This pronoun may function as an expletive anticipatory pronoun holding the syntactic position of embedded clauses (cf. section 4.5.).

Observe from (28) and (29) that weather verb constructions consisting of the predicate only and weather verb constructions with a subject NP have the same meaning. In the phrases of (29), the subject NP bears the quasi-θ-role which is assigned by the weather verb. The null-hypothesis is, therefore, to postulate a null expletive pronoun in the subject position of (28) which absorbs this θ-role. The weather verb constructions with an overt expletive is ruled out by the fact that the demonstrative pronoun *az* 'that' must be assigned a referential θ-role. Note, then, that there is a distributional difference between expletive small *pro* and its overt counterpart *az*. The latter may not appear in the subject position of weather verb constructions.

Empirical evidence for this hypothesis is provided by investigating Binding Principle C effects with these constructions. Consider the following sentences:

(30) a. Esik (*az eső*) csak úgy zuhog (**az eső*)
rains the rain just as pours the rain
'It is pouring with rain.'
b. Esik (*az eső*) csak úgy locsog (**az eső*)
rains the rain just as plashes the rain
'It is plashing with rain.'

In these expressions, the subject of the matrix clause is intended to be coreferential with the subject of the embedded clause. The predicate of both the matrix clause and the embedded clause is a weather verb which may appear independently with an overt NP (cf. (28b)-(28c) and (29b)-(29c)). Note that under the coindexing in (30) the overt NP *az eső* in the embedded clause may not be spelled out.

This fact may be accounted for along the following lines. The NP *az eső* is an R-expression. Hence, its distribution when it is coreferent with another NP is determined by Binding Principle C. If the subjects of both the matrix clause and the embedded clause are overt NPs, *az eső* in the embedded clause may not be spelled out. This is due to the fact that it is bound by the subject NP of the matrix clause. This yields then a Binding Principle C violation. Nor may *az eső* be spelled out in the subject position of the embedded clause when the subject NP of the matrix clause is omitted. In order to account for the ungrammaticality of a coreferential reading in this case, I will hypothesize that an expletive *pro* is present when there is no overt subject present. Under this assumption these sentences display a configuration which is ruled out by Binding Principle C as well.

This parallel distribution between overt NPs and their non-overt counterparts with Binding Principle C resembles the parallel distribution of overt and null pronominals with principles of the binding theory discussed under (*I*) above. The assumption of an expletive *pro* in Hungarian weather verb constructions also explains why a coreferential reading in (30) is possible when *az eső* in the embedded clause is dropped. If its position is occupied by small *pro* no binding theory violation appears. Small *pro,* being a pronoun, is subsumed under Binding Principle B. Embedded *pro* in (30) is free in its governing category, the embedded clause. This provides support for the assumption that null expletive *pro* is present in weather verb construction.

(*ii*) *Coreference of third person pronouns* also indicates that overt pronouns and their non-overt counterparts do not have the same distribution. Kenesei (1985: fn.6). ob-

serves that the nominative third person personal pronouns *ő* 'he/she'and the accusative third person pronoun *őt* 'him/her', can only have [+human] referents. The demonstrative pronoun *az* 'that' refers to [-human] referents. The dropped versions of the nominative and accusative third person personal pronouns, however, may refer both to [+human] and [-human] referents. Compare:

(31) a. Mari látta a könyvet, de nem olvasta *(azt)/(*őt)*
Mary saw-AGR3sg the book-ACC but not read-AGR3sg that-ACC/him
'Mary saw the book, but she didn't read it.'
b. Mari látta a könyvet, de nem írt *(?*arról)/*(róla)*
Mary saw-AGR3sg the book-ACC but not wrote-AGR3sg that-DELAT/it-DELAT
'Mary saw the book but she didn't write about it.'
(Kenesei 1985: 163)

This shows that the coreference with nominative and accusative third person *pro* has a wider range of antecedents than its overt nominative and accusative counterparts.

(*iii*) The *impersonal passive construction* in English is formed by means of the rule of there-insertion. Consider:

(32) *There* is ringing

Hungarian employs a different strategy. The impersonal passive construction is rendered by a *third person plural missing subject construction.* The subject personal pronoun must be dropped. Otherwise the sentence would receive an active interpretation with the pronoun functioning as a referential expression. Compare:

(33) a. Ők csengetnek
they ring-AGR3pl-indef
'They are ringing.'
b. Csengetnek
'There is ringing.'

In accordance with the Extended Projection Principle (cf. 3.3.(7)), I will assume that small *pro* is present in the subject position of (33b) which absorbs the agent role of the verb *csenget* 'to ring'. Clause (33b) may be translated, in fact, more correctly as *someone is ringing.* So, the difference between (33a) and (33b) does not lie in an active-passive dichotomy but rather in that the overt pronoun is *specified,* whereas small *pro* is *unspecified.* The latter yields the impersonal passive construction in Hungarian. This implies that a subjective third person plural *pro* need not have an overt counterpart.

Recapitulating, I have presented two types of arguments in favor of *pro* in the syntax of Hungarian. (*I*) The parallel distribution of overt and their non-overt counterparts with Binding Principles B and C. A non-overt pronoun must be assumed in the position of omitted ones in order to account for the identity of coreference possibilities. (*II*) Overt pronouns and their non-overt counterparts may have a different distribution. Null expletive *pro* may function as the subject in weather verb constructions. Nominative and accusative third person pronouns may only refer to [+human] antecedents, whereas their non-overt counterparts may also corefer with [-human] antecedents. Small *pro* but not an overt third person plural pronoun may be the subject of an impersonal passive construction. This division of functions bet-

ween overt and non-overt personal pronouns provides an argument for the independent status of *pro* in the grammar. Having provided evidence for the presence of this category in the syntax of Hungarian, let us determine its *distribution.*

4.2.4.2. *The Distribution of pro in Hungarian*

The conditions under which personal pronouns can be dropped have been captured in the *Pro-drop Parameter* (cf. Chomsky 1981; among others). Informally, this parameter states that personal pronouns may be omitted in a language if that language possesses "rich" person-number inflection. Theories about the licensing of *pro* rely on the concept of *local recovery*. This involves two subparts, namely the conditions specifying its *structural* sanctioning and conditions specifying its *ϕ-features.* Rizzi (1986), which I will follow here, proposes a theory of licensing conditions of *pro.* The structural sanctioning of *pro* is linked to the presence of a Case-assigning head. This head may belong to a language-specific set, like I[+AGR] in Romance. The feature specification of *pro* is licit only when it is recovered through a binding relation with a head bearing AGR-features. Rizzi further argues that a successful recovery of the person and number features is a necessary condition for functioning as a referential NP.

The phenomenon of *pro*-drop in Hungarian is *more* extensive than in Romance. As we have observed above not only nominative pronouns but also accusative pronouns may be dropped. The phenomenon is further conditioned by the distribution of the *conjugation-type* of the verb. Recall that first and second person accusative pronouns trigger indefinite conjugation, whereas third person accusative pronouns trigger definite conjugation.

Observe from the paradigms in (7) and (8) that nominative personal pronouns may be dropped in all persons and numbers both in the indefinite and definite conjugation. Accusative personal pronouns, on the other hand, may only be dropped in the singular. (This is also the case with the verbal suffix *-lak,* which signals that the nominative subject is first person singular and the accusative object is second person singular or plural (cf. (8a)). Recall that pronominal forms of the lexical cases may not be dropped. Summarizing, *pro*-drop in Hungarian has the following distribution:

(34) *The Distribution of pro in Hungarian*
   a. *Nominative* personal pronouns may be dropped in all persons and numbers
   b. *Accusative* personal pronouns may be dropped only in case they are singular. First and second person pronouns may be dropped with the indefinite conjugation. Third person pronouns may be dropped only with the definite conjugation
   c. Personal pronouns with lexical case may not be dropped

Let us determine how the distribution of *pro* in Hungarian is related to Rizzi's (1986) theory of *pro*-drop.

The question is how *pro* is licensed in Hungarian. *Structurally,* nominative and accusative *pro* may be licensed by I[+AGR] and V respectively, which are both Case-assigning heads (cf. 3.2.(7)). If we assume that the licensing of *pro* is related to structural Case, it is obvious why pronouns with lexical case (cf. (21)) may not be

dropped. Recall that lexical case is thematically governed (cf. section 3.2.1.). This yields the following generalization on *pro*-drop in terms of *Case theory:*

(35) Pronouns in Hungarian may only be dropped if they are assigned structural Case

The licensing of the content of *pro* is connected to the AGR-features on the verbal head.[9] An apparent problem for this hypothesis is the absence of overt AGR in the case of the indefinite conjugation third person singular (cf. (8c)). Note, however, that in this case as well I has *discrete* grammatical features. The gap in the indefinite paradigm is unambiguously marked by absence of all other phonetically represented members of the relevant paradigm. Therefore, the zero-realization in (8c) has exactly the same status as any other realization of AGR.

As may be clear from (34), asymmetries show up between the nominative subject and the accusative object with respect to the licensing of *pro.* The AGR-features of both the definite and the indefinite pattern are "rich" enough to recover the features of non-overt nominative pronouns but obviously cannot license all persons and numbers in the accusative paradigm. If no additional constraints were operative we would end up with *ambiguities* in cases as (7) and (8). However, the outranking of plural by singular in both conjugational patterns, and the prominence of first person singular over the second person singular and plural in the case of the verbal suffix *-lak* suggest that there is an association between the phenomenon of *pro*-drop and *discourse.*

The discourse helps to reduce ambiguities. The restrictions in discourse which condition the "filling in" of the content of *pro* have the form of *individuation hierarchies* (cf. Timberlake 1975). According to Timberlake, individuation is the degree to which the participants are characterized as a distinct entity or individual in discourse. Timberlake proposes the following individuation hierarchies (cf. also Silverstein 1985):

(36) *Individuation hierarchies*
a. 1 > 2 > 3 b. sg > pl

So, first person is higher on the scale than second or third, in the sense that its referent is more highly individuated than second and third person. First and second person are more highly individuated (the speaker and hearer are uniquely referential in the clause) than third person which is assigned reference only in discourse. Singular has a higher degree of individuation than plural.

Therefore, we formulate the following rule which applies at the interface between syntax and discourse:

(37) If structural and morphological conditions do not sanction *pro* unambiguously, then apply *pro*-drop in agreement with the hierarchies in (36)

For example, verbal morphology and structural configuration cannot disambiguate accusative *pro*-drop. The feature *number* of accusative personal pronouns is not

(9) Besides the licensing of *pro* by AGR, Huang (1984) observes that in languages such as Chinese, Japanese, or Korean *pro* may be licensed by an antecedent in discourse. Huang argues that this type of *pro*-drop is a subcase of a more general property of those languages, namely the property of being discourse-oriented.

recoverable. Hence, in accordance with (37) only object singular pronouns may be omitted. I will leave the elaboration of the precise relation between *pro*-drop and discourse strategies as a topic for further research.

Whatever the exact principles are which determine *pro*-drop in Hungarian, the rather specific, not to say bizarre, distribution of *pro* in Hungarian (cf. (33a) and (33b)) provides an excellent diagnostic for "knowing" when there is a small *pro* present in the syntactic representation.

4.2.5. *Summary*

The system of personal pronouns in Hungarian provides two pieces of evidence in favor of the Projection Principle. Firstly, I noted that the nominative and accusative personal pronouns may be dropped. The presence of a pronominal empty category in such cases is provided by the Projection Principle together with the θ-criterion. Evidence from the distribution of overt and omitted pronouns has shown that this is indeed the case and that this pronominal empty category is Chomsky's (1982) small *pro*. Further, I have specified in (34) the distribution of *pro*. The conditions under which pronouns in Hungarian may be omitted depend on structural configurations, verbal AGR-features, and individuation hierarchies in discourse. Secondly, I observed that the stem of personal pronouns with lexical case is often homophonous with the corresponding case-suffix. It follows from this property and the requirement that pronouns with lexical case may not be dropped (cf. (34c)) that an argumental pronominal with lexical case is always visible at surface structure. This is in agreement with the Projection Principle. The Φ-features of the personal pronoun with a case-stem are specified by adding AGR to the case-stem.

4.3. Left Dislocation in Hungarian

Consider the following clauses:

(1) a. *Mari, ő/az* szereti Imrét
Mary she/that loves Imre-ACC
'Mary, she loves Imre.'

b. *Marit, őt/azt* szereti Imre
ACC she-ACC/that-ACC loves Imre
'Mary, Imre loves her.'

c. *Marinak, neki/annak* nem adtam semmit
Mary-DAT she-DAT/that-DAT not gave nothing-ACC
'Mary, I did not give her anything.'

d. *Marival, vele/azzal* találkoztam tegnap
Mary-INSTR she-INSTR/that-INSTR met yesterday
'Mary, I met her yesterday.'

e. *Marira, rá/arra* sokat gondoltam
Mary-SUBL she-SUBL/that-SUBL a lot thought
'Mary, I have thought a lot of her.'

f. *Maritól, tőle/attól* kaptam egy könyvet
Mary-ABL she-ABL/that-ABL got a book-ACC
'Mary, I got a book from her.'

The above clauses are instances of *Left Dislocation* in Hungarian. The left-dislocated NP is pronounced with a rising intonation and is separated from a clause by a pause indicated by a comma in (1).

The pronoun has the following properties. (*i*) It bears stress. (*ii*) The pronoun coreferential with the left-dislocated NP may appear either as a personal or as a demonstrative pronoun. This personal pronoun/demonstrative-switch is subject to dialectal variation. Anna Szabolcsi (personal communication) informs me that in her dialect only the demonstrative pronoun is used. (*iii*) The pronoun must be right-adjacent to the left-dislocated NP, that is, in clause-initial position. (*iv*) It may not be omitted even when it is associated with AGR and satisfies the diagnostics of *pro*-drop (cf. (34)).[10] (*v*) It bears the lexical case assigned by the verb. Note that the verbs *ad* 'give', *találkoz* 'meet', *gondol* 'think', and *kap* 'get' subcategorize for a lexical *dative, instrumental, sublative,* and *ablative* in (1c)-(1f) respectively.

At this place, I will not present an exhaustive analysis of this phenomenon (cf. De Groot 1981b for discussion) but I will rather concentrate on the question why a pronoun is present in the clause.

In the literature on Left Dislocation (cf. Van Riemsdijk and Zwarts 1974; Koster 1987; among others), it has been argued that clauses such as:

(2) *That book*, I won't read *it*

are not derived by an application of move-α. Instead the left-dislocated NP *that book* is *base-generated* outside the clause in a non-A-position which is adjoined to the sentence. The left-dislocated NP depends for its Case- and θ-features on the pronoun with which it is coreferential, in (2) *it*.

Left Dislocation in German indicates that this rule may not only transfer θ-but also Case-features, as has been pointed out in Koster (1987: 65). Consider:

(3) *Den Hans*, ich habe *ihn* gestern gesehen
the John-ACC I have him yesterday seen
'John, I saw him yesterday.'
(Van Riemsdijk 1978: 175)

Folllowing Koster (1987: 65), I will assume that Left Dislocation is non-transformationally derived and has the properties discussed in connection with the clauses (2) and (3). Having settled this, let us return to the Hungarian cases in (1) and provide an answer to the question put forth above.

We observed that the anaphoric pronoun in the sentence satisfies the Case- and θ-requirements of the verb which are transferred to the connected NP in left dislocation position. The presence of the pronoun in the local domain of the verb, i.e. the clause, can only be guaranteed if the Projection Principle is operative which maps lexical requirements onto the overt syntactic representation.

(10) Anna Szabolcsi (personal communication) reports that Counterfocus is an instance of Left Dislocation with *pro* instead of an overt pronoun (cf. Szabolcsi 1981b; 1981c, and Kenesei 1984c for the phonetics and semantics of this construction):

(i) Marit, *pro* SZEreti Imre
Mary-ACC her loves Imre

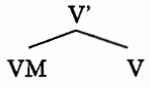
4.4. Complex Verb Constructions in Hungarian

In this section, I discuss the syntactic behavior of lexical items in Hungarian which may function either as *personal pronouns* or as *Verbal Modifiers (VM).* In the former case they have an argumental interpretation, whereas in the latter case they have a non-argumental one. This difference is due to an interaction of lexical properties and the Projection Principle. Before determining how the Projection Principle operates in these cases, let us first discuss *complex verb constructions* in Hungarian.

4.4.1. *The Structure of Complex Verb Constructions*

I noted briefly in chapter two that Hungarian possesses a productive strategy to form *complex verbs*. According to Ackerman and Komlósy (1983), these verbs consist of a VM and a V, and may be represented categorially as V':

(1)



Verbal prefixes may also function as VMs. Ackerman and Komlósy argue that verbal prefixes have no independent θ-role and therefore they treat them as affixes in the sense of Lieber (1980). Such affixes may subcategorize for other morphemes. The lexical entries of affixes indicate both the category of items to which they attach and the category of items produced. The verbal prefix *meg* 'perfectivity marker' has the following subcategorization frame:

(2) *meg:* [V' - [V]]

Context-free rewrite rules and feature percolation conventions guarantee that a complex verb is formed and that it receives a new category label.

Ackerman and Komlósy present the following evidence for this V'-constituency. Firstly, the *word order* of the [VM-V] combination is restricted. In their neutral order, VMs must appear immediately in preverbal position (cf. (3a)). The neutral order is characterized by a level-prosody intonation in the sense of Kálmán et al. (1986). On the other hand, the verbal prefix must be postposed in non-neutral orders, like in (3b) in which the accusative NP *házat* is focussed.[11] Compare:

(3) a. Mari meg vette a házat
Mary perf bought-AGR3sg the house-ACC
'Mari has bought the house.'
b. Mari a házat vette meg
'It was the house that Mary has bought.'

Secondly, the VM and the V may not be separated by sentence adverbs (cf. also Horvath 1981). Therefore, strings with these adverbs (ADV) and [VM-V] combinations pattern in the following manner:

(4) a. ... ADV VM V ... b. ... VM V ADV ... c. *... VM ADV V ...

(11) See for derivation of the inverse-order of the [VM-V] section 2.2.

Consider an example with the sentence adverb *remélhetőleg* 'hopefully'.[12]

(5) a. A fiú *remélhetőleg* be fejezte a feladatot
the boy hopefully perf-finished the assignment-ACC
'Hopefully, the boy has finished the assignment.'
b. A fiú befejezte *remélhetőleg* feladotot
c. *A fiú be *remélhetőleg* fejezte a feladatot

Thirdly, [VM-V] combinations may interact with the *morpholexical* component of the grammar. They may be input to derivational processes which create verbs and nominals from a [VM-V] sequence. Consider for example the following derivations:

(6) a. *győz* 'win'
b. *meggyőz* 'convince'
c. **győződik*
d. *meggyőződik* 'be convinced of'

The verb *győz* in (6a) may be prefixed with the VM *meg* 'perfectivity marker' deriving the verb *meggyőz* (cf. (6b)). As may be observed from the ungrammaticality of (6c), *győz* cannot be suffixed with the verbal derivational suffix *-ődik*, a passivizer with the properties in 3.3.(10)). This suffix relates for example the verbs *nyel* 'swallow', and *nyelődik* 'is swallowed'. The derived lexical item *meggyőződik* in (6d), however, contains both the prefix *meg* and the suffix *-ődik*.

The question is now how it is derived? The input to this verbal item cannot be (6c) because *győz* to which *-ődik* is attached is a lexical gap. Another possibility is that (6d) is formed by attaching *-ődik* to the verb *meggyőz* in (6b). If this derivation takes place in the lexicon, then the prefix *meg* must be attached to the verb *győz* already in the lexicon.

Fourthly, [VM-V] combinations may have *different* substructures in their Predicate Argument Structure (PAS) than the basic verb which participates in the complex verb construction. The reason for this difference is that the attachment of VMs may affect the substructures in the lexical entry of a basic verb. The fact that the PAS of complex verb constructions is already determined in the morphological component of the lexicon demonstrates that they are already merged into a V' in this component of the grammar. For example, the verb *tol* 'push' is a tryadic verb subcategorizing for a NOM-ACC-ILL case frame (cf. (7a)). However, when combined with the verbal prefix *meg* 'perfectivity marker' it turns into a dyadic verb with a NOM-ACC case frame (cf. (7b)). Compare:

(7) a. János a sarokba tolta a szekrényt
John the corner-ILL pushed the cupboard-ACC
'John pushed the cupboard in the corner.'
b. János *meg* tolta a szekrényt
John perf-pushed the cupboard-ACC
'John has pushed the cupboard.'
c. *János *meg* tolta a szekrényt a sarokba

(12) The fact that sentence adverbs may be interposed between the verb and its accusative object will be discussed in section 5.2.1.1. and 5.4.

4.4.2. *Complex Verb Constructions and the Projection Principle*

This section discusses complex verb constructions consisting of a *verbal prefix* and a verb. I will focus on the following four prefixes:

(8) a. *neki* '(in)to
b. *bele* 'into'
c. *rá* 'on'
d. *hozzá* 'to'

The reason we isolated these prefixes is that they are *homophonous* with the dative, illative, sublative, and allative personal pronouns of the third person singular. Compare:

(9) a. *neki*
DAT-AGR3sg
'to him/her'
b. *bele*
ILL-AGR3sg
'into him/her'
c. *rá*
SUBL-AGR3sg
'on(to) him/her'
d. *hozzá*
ALL-AGR3sg
'to him/her'

The *double-faced* nature of these items offers an excellent opportunity for providing insight into the application of the Projection Principle in Hungarian.

The verbal prefixes in (8) may select a verb of the semantic class of *linear motion.* Compare, for example, the lexical entry of the Hungarian verb *fut* 'run' which contains the following substructures, among others:

(10) LCS for Hungarian 'run': {x moves along a path rapidly}
θ-grid for Hungarian 'run': (agent)

Because the variable *x* undergoes a change of location it could also be defined as the theme of the action denoted by the verb (cf. Hale and Laughren 1983). However, whatever the exact θ-role is of the argument selected by the verb, it always appears in the nominative case:

(11) Mari futott
Mary ran-AGR3sg
'Mary has run'

Attaching the prefixes in (8) to the verb *fut* 'run' has the following consequences for the substructures of its lexical entry. In the LCS an entity is added corresponding to the place into/to/at/on which the agent is going. This entity is associated with the *goal.* From this it follows that these VMs introduce an *extra* argument. Therefore, these prefixes function as argument taking predicates (ATP). The argument added is assigned dative, illative, sublative, or ablative depending on whether *neki, bele, rá,* or *hozzá* is prefixed respectively to the verbal stem *fut* 'run'. Further, LCS indicates how the goal is affected by the agent. Compare some of the substructures of the lexical entry of the complex verbs formed by the verb *fut* and these prefixes:

(12) a. *nekifut:*
LCS for *nekifut:* {*x* moves along a path rapidly toward y such that it comes into contact with *y*}
θ-grid for *nekifut:* (agent, goal)
case frame for *nekifut:* NOM run into DAT

b. *belefut:*
LCS for *belefut:* {*x* moves along a path rapidly toward y such that it comes to be internal to *y*}
θ-grid for *belefut:* (agent, goal)
case frame for *belefut:* NOM run into ILL

c. *ráfut:*
LCS for *ráfut:* {*x* moves along a path rapidly toward y such that it gets on the surface of *y*}
θ-grid for *ráfut:* (agent, goal)
case frame for *ráfut:* NOM run on SUBL

d. *hozzáfut:*
LCS for *hozzáfut:* {*x* moves along a path rapidly toward y such that it comes into facinity to *y*}
θ-grid for *hozzáfut:* (agent, goal)
case frame for *hozzáfut:* NOM run to SUBL

Observe the following sentences with the verb *ráfut* (cf. (12c)). This choice does not affect the course of the argumentation below. In fact, examples with any of these verbs could have been chosen. Compare:

(13) a. *Rá* [ATP] futott a hegyre [ARG]
onto ran-AGR3sg the mountain-SUBL
'He ran onto the mountain.'

b. **Rá* [ATP] futott
onto ran-AGR3sg

c. **Rá* [ARG] futott a hegyre [ARG]
it-SUBL ran-AGR3sg the mountain-SUBL

d. **Rám* [ARG] futott a hegyre [ARG]
I-SUBL ran-AGR3sg the mountain-SUBL

e. *Rá* [ARG] futott
it-SUBL ran-AGR3sg
'He ran onto it.'

f. *A hegyre* [ARG] futott
the mountain-SUBL ran-AGR3sg
'He ran onto the mountain.'

g. *Rám* [ARG] futott
I-SUBL ran-AGR3sg
'He ran onto me.'

As already noted, the prefixes in (8) are homophonous with the dative, illative, sublative, and allative personal pronouns of the third person singular. Further, we noticed that these prefixes may function as ATPs, whereas they may be *argumental* (ARG) as personal pronouns.

In (13a), *rá* 'onto' functions as a prefix and the NP a *hegy* 'the mountain' is associated with the sublative argument in the case frame of *ráfut.* The ungrammaticality of (13b) shows that the sublative argument may not be omitted. The sentences in (13c) and (13d) exemplify that *rá* 'it-SUBL' and *rám* 'I-SUBL' respectively may not receive an argumental interpretation when another sublative argument, i.e. a *hegyre* 'the mountain-SUBL', is present in the sentence. The reason for the ungrammaticality in (13c) and (13d) is not caused by the absence of a verbal prefix, for the sentences in (13e)-(13g) demonstrate that the verb *fut* 'run' may always surface with an optional sublative NP which receives an argumental interpretation.

From this paradigm, we may draw the following conclusions. The comparison of (13a) with (13b) demonstrates that the mapping of LS onto phrase structure is *obligatory.* Argumental NPs, selected, may not be omitted. Further, the sentences (13a) versus (13c) or (13d) show that the relation between LS and phrase structure is also

restricted by a biuniqueness condition. In case *rá* or one of its inflected alternants and a full referential NP are present, the sentence receives only a grammatical reading if it is possible to interpret *rá* as a verbal prefix, such as in (13a). In (13d), this is impossible because *rá* is inflected for the first person singular. Therefore, (13d) has no grammatical counterpart.

4.5. Embedded Clause Formation in Hungarian

In this section, I will discuss the formation of embedded *clauses* in Hungarian. I will conclude that their shape supports the hypothesis that the Projection Principle is operative in Hungarian. Before discussing some *linear* restrictions on the formation of embedded clauses, let us first turn to a discussion of their *structure.*[13]

4.5.1. *The Structure of Embedded Clauses*

Hungarian distinguishes two types of *subordination.* Embedded clauses may either be related to a constituent of the matrix sentence, or may appear freely in the matrix sentence.[14] In the present context only a discussion of the former type is relevant.

Kenesei (1985) observes that the NPs to which embedded clauses are related may be of two types: they are either *lexical* or *pronominal* ('anticipatory'). This pronoun is homophonous with the non-proximate demonstrative pronoun *az* 'that', or with the third person singular personal pronoun. In this section only examples with the demonstrative anticipatory pronoun will be presented, postponing the discussion of embedded clauses related to a personal anticipatory pronoun until the following section. These two types of constituents may be used in the formation of both *relative* and *that-clauses* in Hungarian. Compare:

(1) a. *Relative, lexical NP*
Az a darab, amit Péter látott, érdekes volt
that the play what-ACC Peter saw-AGR3sg interesting was
The play that Peter saw was interesting.'

b. *Relative, anticipatory pronoun*
Az, amit Péter látott, érdekes volt
that what-ACC Peter saw-AGR3sg interesting was
'That what Peter saw was interesting.'
(Kenesei 1985f: 145)

(2) a. *That-clause, lexical NP*
Az a kérdés, hogy mit látott Péter, érdekes
that the question that what-ACC saw-AGR3sg Peter interesting
'The question of what Peter saw is interesting.'

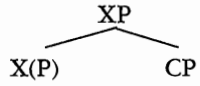
(13) Anna Szabolcsi (personal communication) brings to my attention that there is an alternant of the sublative, illative, and allative third person singular pronoun which dissolves the syntactic ambiguity between the verbal prefix and personal pronoun function. By adding the suffix *-ja/je* of the third person possessive agreement to *rá, bele,* and *hozzá,* they are turned unambiguously into personal pronouns: *rája* 'on him/her', *beléje* 'into him/her', and *hozzája* 'to him/her'.

(14) To the latter type belong embedded clauses introduced by complementizers such as *mivel* 'since', *bár* 'though', and free relatives. (See Kenesel 1985a, 1985f and section 5.4. for these cases of subordination with pronominal noncoreference).

b. *That-clause, anticipatory pronoun*
Az, hogy mit látott Péter, érdekes
that that what-ACC saw-AGR3sg Peter interesting
'What Peter saw is interesting.'
(Kenesei 1985f: 146)

Two theories on the structure of embedded clauses are possible.

(*I*) Kenesei (1984a) assumes that embedded clauses of the above type have the following structure:

(3)


The head of this structure is the *X(P)* in which *X* may be substituted by *N, A*, or *P.* In the embedded clauses (1) and (2), the position of (X)P is either filled by a lexical NP or by an anticipatory pronoun. Both constituents are categorially of the type N. This implies that 'ordinary' that-clauses in Hungarian are complex NPs under this hypothesis.

(*II*) A second analysis of embedded clauses in Hungarian relies on the syntactic position embedded clauses may occupy. In general, embedded clauses cannot be in a Case-position, because of the *Case Resistance Principle* (cf. Stowell 1981). This principle states:

(4) *Case Resistance Principle* (CRP)
Case may not be assigned to a category bearing Case-assigning features

Stowell assumes that the feature-matrix of CP contains the feature [+Tense]. This feature is a a Case-assigning feature (cf. Chomsky 1981). Hence, CPs cannot be in a Case-position but must be dislocated.

The question arises what the role of the anticipatory pronoun is under this hypothesis. Compare some other examples with *that*-clauses in which the anticipatory pronoun appears:[15]

(5) a. Kiderült (*az*) [CP hogy János nem olvas]
out-turned-AGR3sg-indef that that John not read-AGR3sg
'It has turned out that John doesn't read.'
b. Tudom (*azt*) [CP hogy János nem olvas]
know-AGR1sg-def that-ACC that John not read-AGR3sg
'I know that John doesn't read.'
c. Hiszek *(*abban*) [CP hogy újra találkozni] fogunk
believe-AGR1sg-indef that-INESS that again meet-INFI will-AGR1pl
'I believe that we will meet again.'
d. Számitok *(*arra*) [CP hogy Mari beteg lesz]
count-AGR1sg-def that-SUBL that Mary ill will-be
'I expect that Mary will be ill.'
e. Tudok *(*arról*) [CP hogy János nem olvas]
know-AGR1sg-indef that-DELAT that John not read-AGR3sg
'I know that John does not read.'

(15) Embedded clauses are inherently definite. In (5b), the embedded clause is associated with the accusative position in the LS of the verb. Hence, the definite conjugation on ther verb.

f. Péter haragszik *(*azért*) [CP hogy Mari megérkezett]
Peter be angry-AGR3sg-indef that-CAUS that Mary arrived-AGR3sg
'Peter is angry because Mary arrived.'

In (5a)-(5f), the verbs *kiderül* 'turn out', *tud* 'know', *hisz* 'believe', *számít* 'count on', *tud* 'know about', and *haragszik* 'be angry' appear with a NOM, NOM-ACC, NOM-INESS, NOM-SUBL, NOM-DELAT, and NOM-CAUS case frame. If embedded clauses are in a non-A-position, as we hypothesized above, then the Case- and θ-features of the verb cannot be satisfied by the CP. Therefore, they are absorbed by the anticipatory "dummy" pronoun *az,* which is base-generated in an A-position and linked to the CP.[16] This covers the fact that the above anticipatory pronouns bear nominative, accusative, inessive, sublative, delative, or causalis case.

It seems to me that the analysis of embedded clauses in (II) should be preferred over the one in (I), because it is related to general principles of the grammar. As a consequence, the function of *az* is comparable to the function of English *it* and Dutch *het* with a postverbal CP. Compare:

(6) a. Ik betreur *het* [CP dat Jan ziek is] b. *It* surprised me [CP that John is ill]
I regret it that John ill is

Bennis (1986: ch.2) and Koster (1987: ch.5) argue that *het* and *it* are referential expressions in an A-position carrying a propositional θ-role. This analysis accounts for the fact that extraction may not take place from postverbal embedded clauses, since they are in adjunct position. Adjunct clauses usually form islands for extraction. Hence, the ungrammaticality of the following sentence:

(7) *Wat betreurde jij het [CP dat hij had gezegd]?
what regretted you it that he had said

If CPs in Hungarian are in an adjunct position and their position in the LS of the verb is occupied by *az,* then we expect that extraction from embedded clauses will be ruled out. In chapter six, I will demonstrate that this is indeed the case. Therefore, this provides further support for the analysis adopted here.

Before I discuss some *linear* restrictions on the position of embedded clauses in Hungarian, let us first consider the omissibility of anticipatory pronouns in (5).

The nominative and accusative anticipatory pronoun may be dropped (cf. (5a), (5b)) but not the anticipatory pronouns with lexical case (cf. (5c)-(5f)). Note that this corresponds with the distribution of small *pro* in Hungarian (cf. 4.2.(34)). Therefore, I will assume that the anticipatory pronoun is replaced by *pro* when omitted (cf. also Kenesei 1984a; 1985d for a similar claim). This implies that there are no free that-clauses in Hungarian.

(16) There are a number of anticipatory pronoun-complementizer pairs which introduce adjunct embedded clauses, like embedded clauses of time such as the pairs *akkor-amikor* 'then-when' and *azalatt-mialatt* 'during it-while', or embedded clauses of place such as *ott-ahol* 'there-where', etc. Compare:

(i) a. *Azalatt mialatt* János keményen tanult Péter lányokhoz járt
it-during while John hard studied Peter girls-ALL went
'While John was studying hard, Peter went to meet girls.'
b. *Ott ahol* sok a titkos rendör nem jó lakni
there where lot the secret agent not good live-INFI
'Where a lot of secret agents are, it is not good living.'

4.5.2. *Linear Restrictions on Embedded Clauses*

In the preceding section, I examined the structure of embedded clauses in Hungarian. Let us turn now to a discussion of *linear* restrictions on their position in the sentence. Although embedded clauses may be scrambled around freely, the following linear restrictions may be observed:

(*I*) Kenesei (1984a) observes that *that-clauses* and the lexical NP or the anticipatory pronoun to which they are related may be separated by intervening material. The sentences in (8) are the contiguous counterparts of the sentences in (2) (bracketing is mine):

(8) a. *Az a kérdés* érdekes [$_{CP}$ hogy Péter mit látott]
that the question interesting that Peter what-ACC saw-AGR3sg
b. *Az érdekes* [$_{CP}$ hogy Péter mit látott]
that interesting that Peter what-ACC saw-AGR3sg

Kenesei notes furthermore that the CP and its related lexical NP or anticipatory pronoun must be *non-adjacent* obligatorily when the embedded clause is focussed.[17] Compare:

(9) a. *[$_{F}$ *Az a kérdes* [$_{CP}$ hogy Péter mit látott]] érdekes
b. [$_{F}$ *Az a kérdés*] érdekes [$_{CP}$ hogy Péter mit látott]
'It is the question what Peter saw that is interesting.'
c. *[$_{F}$ *Az* [$_{CP}$ hogy Péter mit latott]] érdekes
d. [$_{F}$ *Az*] érdekes [$_{CP}$ hogy Péter mit látott]
'What Peter saw IS interesting.'

In accordance with the analysis of embedded clauses argued for above, the lexical NP or the anticipatory pronoun is in the Focus position, and the CP is base-generated postverbally.

(*II*) Another linear restriction on embedded clauses has been discussed in Kenesei (1984a; 1985d). According to Kenesei, there are positional restrictions on the occurrence of the anticipatory pronoun. If the order '*that*-clause...V... demonstrative pronoun' occurs, the third person singular personal pronoun must replace its corresponding demonstrative anticipatory pronoun. This pronoun always has the same Case-marking as the anticipatory pronoun. Compare the scrambled variants of the sentences in (5):

Demonstrative/personal pronoun - that-clause - V

(10) a. **(Az)/*ő* [hogy János nem olvas] kiderült
that/he that John not read-AGR3sg out-turned-AGR3sg
b. **(Azt)/*őt* [hogy János nem olvas] tudom
that-ACC/he-ACC that John not read-AGR3sg know-AGR1sg
c. *Abban/*benne* [hogy újra találkozni fogunk] hiszek
that-INESS/it-INESS that again meet-INFI will-AGR1pl believe-AGR1sg
d. *Arra/*rá* [hogy Mari beteg lesz számítok]
that-SUBL/it-SUBL that Mary ill will-be count-AGR1sg

(17) Compare for discussion of Extraposition É. Kiss (1981a) and Kenesei (1984a). Compare, furthermore, Kenesei (1985e) for the interaction of constituent embedding and the uniformity condition on the branching of X'-categories (cf. 2.2.1.(1)) yielding Extraposition obligatorily.

e. *Arról/*róla* [hogy János nem olvas] tudok
that-DELAT/it-DELAT that John not read-AGR3sg know-AGR1sg
f. *Azért/*érte* [hogy Mari megérkezett] Péter haragszik
that-CAUS/it-CAUS that Mary arrived-AGR3sg Peter is-angry

That-clause - V - demonstrative/personal pronoun

(11) a. [Hogy János nem olvas] *(*az*)/*ő* kiderült
that John not read-AGR3sg that/he out-turned-AGR3sg
b. [Hogy János nem olvas] *(*azt*)/*őt* tudom
that John not read-AGR3sg that-ACC/he-ACC know-AGR1sg
c. [Hogy újra találkozni fogunk] *abban/*benne* hiszek
that again meet-INFI will-AGR1pl that-INESS/it-INESS believe-AGR1sg
d. [Hogy Mari beteg lesz] *arra/*rá* számítok
that Mary ill will-be that-SUBL/it-SUBL count-AGR1sg
e. [Hogy János nem olvas] *arról/*róla* tudok
that John not read-AGR3sg that-DELAT/it-DELAT know-AGR1sg
f. [Hogy Mari megérkezett] *azért/*érte* Péter haragszik
that Mary arrived-AGR3sg that-CAUS/it-CAUS Peter is-angry

Demonstrative/personal pronoun - V - that-clause

(12) a. *(*Az*)/*ő* kiderült [hogy János nem olvas]
that/he out-turned-AGR3sg that John not read-AGR3sg
b. *(*Azt*)/*őt tudom [hogy János nem olvas]
that-ACC/he-ACC know-AGR1sg that John not read-AGR3sg
c *Abban/*benne* hiszek [hogy újra találkorni fogunk]
that-INESS/it-INESS believe-AGR1sg that again meet-INFI will-AGR1pl
d. *Arra/*rá* számítok [hogy Mari beteg lesz]
that-SUBL/it-SUBL count-AGR1pl that Mary ill will-be
e. *Arról/*róla* tudok [hogy János nem olvas]
that-DELAT/it-DELAT know-AGR1sg that John not read-AGR3sg
f. *Azért/*érte* Péter haragszik [hogy Mari megérkezett]
that-CAUS/it-CAUS Peter is-angry that Mary arrived-AGR3sg

That-clause - V - demonstrative/personal pronoun

(13) a. [Hogy János nem olvas] kiderült *(*az*)/(*ő*)
that John not read-AGR3sg out-turned-AGR3sg that/he
b. [Hogy János nem olvas] tudom *(*azt*)/(*őt*)
that John not read-AGR3sg know-AGR1sg that-ACC/he-ACC
c. [Hogy újra találkozni fogunk] hiszek **abban/benne*
that again meet-INFI will-AGR1pl believe-AGR1sg that-INESS/it-INESS
d. [Hogy Mari beteg lesz] számítok **arra/rá*
that Mary ill will-be count-AGR1sg that-SUBL/it-SUBL
e. [Hogy János nem olvas] tudok **arról/róla*
that John not read-AGR3sg know-AGR1sg that-DELAT/it-DELAT
f. [Hogy Mari megérkezett] Péter haragszik **azért/érte*
that Mary arrived-AGR3sg Peter is-angry that-CAUS/it-CAUS

V - demonstrative/personal pronoun - that-clause

(14) a. Kiderült *(az)/(*ő*)* [hogy János nem olvas]
out-turned-AGR3sg that/he that John not read-AGR3sg
b. Tudom *(azt)/(*őt*)* [hogy János nem olvas]
know-AGR1sg that-ACC/he-ACC that John not read-AGR3sg

c. Hiszek *abban/benne* [hogy újra találkozni fogunk]
believe-AGR1sg that-INESS/it-INESS that again meet-INFI will-AGR1pl
d. Számitok *arra/rá* [hogy Mari beteg lesz]
count-AGR1sg that-SUBL/it-SUBL that Mary ill will-be
e. Tudok *arról/róla* [hogy János nem olvas]
know-AGR1sg that-DELAT/it-DELAT that John not read-AGR3sg
f. Péter haragszik *azért/érte* [hogy Mari megérkezett]
Peter is-angry that-CAUS/it-CAUS that Mary arrived-AGR3sg

V - that-clause - demonstrative/personal pronoun

(15) a. *Kiderült [hogy János nem olvas] *az/ő*
out-turned-AGR3sg that John not read-AGR3sg that/he
b. *Tudom [hogy János nem olvas] *(azt)/(őt)*
know-AGR1sg that John not read-AGR3sg that-ACC/he-ACC
c. *Hiszek [hogy újra találkozni fogunk] *abban/benne*
believe-AGR1sg that again meet-INFI will-AGR1pl that-INESS/it-INESS
d. *Számitok [hogy Mari beteg lesz] *arra/rá*
count-AGR1sg that Mary ill will-be that-SUBL/it-SUBL
e. *Tudok [hogy János nem olvas] *arról/róla*
believe-AGR1sg that John not read-AGR3sg that-DELAT/it-DELAT
f. *Péter haragszik [hogy Mari megérkezett] *azért/érte*
Peter is-angry that Mary arrived-AGR3sg that-CAUS/it-CAUS

The paradigm (10)-(15) has the following properties:

(*i*) If the linear order is '*that*-clause...V...anticipatory pronoun' the demonstrative anticipatory pronoun must be replaced by its corresponding third person singular personal pronoun (cf. (13)). This switch is optional when both the *that*-clause and the demonstrative pronoun are postverbal (except for the nominative and accusative anticipatory pronoun) (cf. (14)). Note, however, that in such cases the pronoun may not be in final-position (cf. (15)). Note, further, that only the demonstrative pronoun is allowed preverbally (cf. (10)-(12)).

The following questions arise in connection with this demonstrative/personal pronoun-switch: What is the reason for this phenomenon and what is the status of the personal pronoun in case it replaces the demonstrative pronoun? Kenesei (1984a; 1985d) suggests that the reason for this pronoun-switch has probably to do with a general condition on anaphora in Hungarian. The linking between the demonstrative anticipatory pronoun *az* and an embedded clause may be understood as an anaphoric relation. Obviously, backward anaphora with a demonstrative pronoun is blocked in the '*that*-clause...V...demonstrative pronoun' order (cf. (11)-(13)). Consequently, only a personal pronoun may be related to a clausal antecedent in those cases. It must be admitted, however, that the conditions governing this pronoun-switch need further investigation.[18]

(18) Kenesei (1985a) reports some exceptions to this phenomenon of pronoun-switch.

(i) If the anticipatory pronoun has no corresponding pronominal form with lexical case such as in the case of translative, essive, formalis, and terminative (cf. section 4.2.3.), the anticipatory pronoun may not be replaced. This yields an ungrammatical variant in case the demonstrative pronoun-personal pronoun switch is obligatory, that is, in the order '*that*-clause...V...anticipatory pronoun'. Compare:

The personal pronoun has the same syntactic status as the demonstrative anticipatory pronoun in (13)-(14). It represents the Case- and θ-features of the embedded clause in the LS of the main verb. First, observe that it does not function as a deictic expression with independent reference. Second, consider the following sentence:

(16) a. El akarok menni *azért/*érte* [$_{CP}$ hogy láthassalak]
away want-AGR1sg go-INFI that-CAUS/it-CAUS that see-SUBJ-AGR1sg2sg
'I want to go in order to see you.'
b. [$_{CP}$ Hogy láthassalak] el akarok menni **azért/*érte*
that see-SUBJ-AGR1sg2sg away want-AGR1sg go-INFI go-INFI
that-CAUS/it-CAUS

The case-marker on the anticipatory pronoun, i.e. the causalis, is not selected by the main verb complex. Hence, the embeddedd clause is not a *direct* argument of this complex. Observe that the phenomenon of pronoun-switch is not possible in (16). Not even in the context *'that*-clause... V...anticipatory pronoun' in which regularly this phenomenon is obligatory (cf. (13)). This implies that the personal anticipatory pronoun is base-generated in an NP-position to where a dislocated clause may be linked. Formally, this relation may be expressed by coindexing the agreement marker of the personal pronoun with the embedded clause. The Case- and θ-features of

---

(i) a. Eljutottam *addig,* [$_{CP}$ hogy engedélyt kaptam]
reached-AGR1sg that-TERM that permission got-AGR1sg
'I reached to get permission.'
b. * [$_{CP}$ Hogy engedélyt kaptam] eljutottam *addig*

(II) The pronoun-switch with the superessive case sometimes behaves irregularly. It may not apply in a context where this phenomenon is usually allowed, for example, in the order 'V...anticipatory pronoun... *that*-clause'. Compare an example with the verb *aggódik* 'worry about' which subcategorizes for a superessive argument:

(ii) Aggódtam *azon/*rajta* [$_{CP}$ hogy Mari beteg volt]
worried-AGR1sg that-SUPER/it-SUPER that Mari ill was
'I was worried about the fact that Mary was ill.'

In some idiomatic expressions the demonstrative/personal pronoun-switch may even take place preverbally with the superessive:

(iii) *Azon/raijta* leszek [$_{CP}$ hogy ...
that-SUPER/it-SUPER be-AGR1g that
'I will do my best to ...'

(III) When a complex verb construction contains a prefix that is homophonous with a personal pronoun bearing lexical case, that is, with the lexical items in 4.4.(8), then the demonstrative anticipatory pronoun may not be replaced by a personal pronoun. Compare an example with the verbal prefix *bele* 'into' which is homophonous with the third person singular illative pronoun *bele* 'into it'.

(iv) a. János bele ment *abba/*bele* [$_{CP}$ hogy eljöjjön]
John into went that-ILL/it-ILL that come-SUBJ-AGR3sg
'John consented in coming.'
b. [$_{CP}$ Hogy eljöjjön] János bele ment **abba/*bele*

If the prefix in (iva) is postposed from its preverbal position because some other constituent is focussed, it is better to omit the demonstrative anticipatory pronoun entirely. Compare:

(v) [$_{F}$ János] ment bele (abba) [$_{CP}$ hogy eljöjjön]

I will assume that a small *pro* is present in the syntactic representation if the demonstrative anticipatory pronoun is absent. This *pro* is then sanctioned by the verbal prefix.

The sentences in (iv) and (v) support the conclusion which we reached in section 4.4.2., namely, that the syntax behavior of the double-faced lexical items is determined by the Projection Principle. If *bele* would be present twice in these sentences, it would be impossible to decide which one is the verbal argument.

the verb may be shared under this coindexing by the personal pronoun and the embedded clause.

(*ii*) Note that the overt nominative and accusative personal pronouns may not participate in the pronoun-switch (cf. (10)-(11)), only their non-overt *pro* counterparts. This reason for this is, as pointed out in Kenesei (1985d), that the overt nominative and accusative personal pronouns may have only [+human] referents (cf. also section 4.2.4.2.).

(*iii*) In the preceding section, I noted that the demonstrative anticipatory pronoun may be dropped in accordance with the distribution of *pro* in Hungarian (cf. 4.2.(34)). There are, however, two apparent exceptions to this generalization with anticipatory pronouns.

(*A*) The nominative and accusative anticipatory pronoun may never be dropped in preverbal position (cf. (10)-(13)). This has probably to do with the fact that pragmatic functions such as Topic and Focus are assigned preverbally. These functions are marked phonetically (cf. section 2.1.). It is reasonable to suppose that phonetic markers may only be assigned to overt elements.

(*B*) Kenesei (1985b: fn.7) notes that some verbal and adjectival predicates allow pro-drop with a demonstrative/personal anticipatory pronoun bearing lexical case. This would constitute a counterexample to generalization 4.2.(34c), which states that pronouns with lexical case may not be dropped. Such predicates include, among others, *örülök* 'be happy' subcategorizing for a dative argument, *kíváncsi lenni* 'be curious about' subcategorizing for a sublative argument, *fél* 'be afraid of' which subcategorizes for an ablative argument, and *kezeskedik* 'be sure of' that subcategorizes for a causalis argument. Compare, for example:

(17) Örülök (*annak*)/(*neki*) [$_{CP}$ hogy jöttél]
be happy-AGR1sg that-DAT/it-DAT that came-AGR2sg
'I am happy that you came.'

It is not clear why these predicates permit a violation of 4.2.(34c). Note that semantically they belong to the same category. These predicates express an emotive state. Maybe this is worth exploring further.

4.5.3. *Summary*

Recapitulating, in this section I have presented empirical support from the formation of embedded clauses in Hungarian for the hypothesis that the Projection Principle maps lexical information onto phrase structure in a one-to-one fashion. Embedded clauses may not appear in an A-position, because of the CRP. I have argued that despite this, Case- and θ-features assigned to embedded clauses by a verbal predicator are always represented in the overt syntactic representation. These lexical properties may be carried by a demonstrative anticipatory pronoun, or its personal pronoun alternant. The switch between a demonstrative anticipatory pronoun and a personal anticipatory pronoun seems to be determined by a sort of anaphoric process, which requires further investigation.

4.6. Split Constituents in Hungarian

The Projection Principle specifies a one-to-one correspondence between LS and syntactic representations. For each argument selected at LS there is a corresponding constituent present in syntax. In this section, I will focus on *split constituents* in Hungarian. This phenomenon apparently violates the one-to-one matching between LS and syntactic representation. I will demonstrate, however, that this is not the case. Split constituents in Hungarian are conditioned by highly specific syntactic and semantic restrictions.

Syntactically, the parts of split constituents involve a predication relation signaled by identity of morphological features such as case, number and so on. Semantically, the parts of split constituents express simple conjunction. These restrictions show that split constituents are rather marked. They cannot appear freely. This is in accordance with the Projection Principle.

Split constituents constitute a subcase of noun modification. In section 4.6.1., I will first discuss the syntax of noun modification. Section 4.6.2. examines its semantics. In section 4.6.3., I will present an analysis of split constituents which is in correspondence with the Projection Principle. Finally, in section 4.6.4., I will investigate split constituents appearing in other languages, such as Warlpiri and German, and conclude that this phenomenon favours a representational approach to grammar over a derivational one.

4.6.1. *The Syntax of Noun Modification*

Roughly, modifier noun constructions may appear in two patterns in Hungarian. Either the combination of the modifier and noun forms a *single constituent* (cf. (1a), (2a)) or the parts may be separated resulting in a so-called *split constituent* (cf. (1b), (1c)) and (2b), (2c)):

(1) a. Mari (a) két biciklit (látta)/látott
Mary (the) two bike-ACC saw-AGR3sg-def/indef
'Mary saw (the) two bikes.'

b. Mari *biciklit* látott *kettőt*
Mary bike-ACC saw-AGR3sg two-ACC
'What Mary saw two of were bikes.'

c. Mari *biciklit* látott, *kettőt*
Mary bike-ACC saw-AGR3sg two-ACC
'Mary saw only bikes and there were two of them.'

(2) a. Mari (a) nagy biciklit (látta)/látott
Mary (the) big bike-ACC saw-AGR3sg-def/indef
'Mary saw (the) big bikes/(bike).'

b. Mari *biciklit* látott *nagyot*
Mary bike-ACC saw-AGR3sg big-ACC
'Mary saw bikes such that they were big.'

c. Mari *biciklit* látott, *nagyot*
Mary bike-ACC saw-AGR3sg big-ACC
'Mary saw only bikes and they were big/big ones.'

Superficially, the variants in (1) and (2) have similar properties. On closer investigation, however, it turns out that there are subtle syntactic and semantic differences between them. Because of the fact that intuitively these variants are "connected" a linguistic approach which unifies them seems to be justified. In section 4.6.3., I will consider two such analyses. Let us first discuss the syntactic properties of the above constructions.

(*I*) As appears from (1a) and (2a) attributive modifiers in single NPs are on a *left* branch in Hungarian, whereas in split constituents the modifier is *separated* from the head noun, see ((1b), (1c)) and ((2b), (2c)). Thus the modifier in single NPs is usually in construction with its head. The parts of split constituents, however, may be scrambled around in the sentence[19] freely.

(*II*) Modifiers in single NPs are *uninflected* (cf. (1a), (2a)). Modifiers in split constituents, on the other hand, are *case-marked* (cf. (1b), (1c), (2b) and (2c)). Note that there is congruence between the case of the head noun and the modifier.

(*III*) Single NPs may be *modified* by a determiner (cf. (1a), (2a)). The parts of split constituents, however, must be bare:

(3) a. **A/egy* biciklit láttam kettőt
the/a bike-ACC saw-AGR1sg two-ACC
b. *Biciklit láttam *a/egy* nagyot
bike-ACC saw-AGR1sg the/a big-ACC
c. **A/egy* biciklit láttam *a/egy* nagyot
the/a bike-ACC saw-AGR1sg the/a big-ACC

(*IV*) If modifiers expressing quantity, like numerals or quantifiers, are in construction with the head noun they always require this head to be in the *singular.* Therefore, (4a) is grammatical, unlike (4b). This restriction does not have to be obeyed with split constituents. The head noun may sometimes appear in the plural as well (cf. (4c), (4d)):

(4) a. Láttam két nyulat
saw-AGR1sg two rabbit-sg-ACC
'I saw two rabbits.'
b. *Láttam két nyulakat
saw-AGR1sg two rabbit-pl-ACC
c. Nyulat láttam kettőt
rabbit-sg-ACC saw-AGR1sg two-ACC
d. Nyulakat láttam kettőt
rabbit-pl-ACC saw-AGR1sg two-ACC

If a plural marker is attached to a non-numeral modifier with split constituents, then the head noun must be in plural too:

(5) a. Láttam nagy biciklit
saw-AGR1sg big bike-sg-ACC
'I saw a big bike.'
b. Láttam nagy bicikliket
saw-AGR1sg big bike-pl-ACC
'I saw big bikes.'
c. *Biciklit láttam nagyokat
bike-sg-ACC saw-AGR1sg big-pl-ACC
d. Bicikliket láttam nagyokat
bike-pl-ACC saw-AGR1sg big-pl-ACC

Observe from the comparison between the pairs in ((4c), (4d)) and ((5c), (5d)) that there must be full morphological concord between the parts of split cons-

(19) There is some uncertainty among native-speakers whether the singular count noun *bicikli* 'bike' in (1b) and (1c) is in F-position, in the preverbal modifier position, or may be in both positions. Here I will follow Szabolcsi (1983c) who presents only examples in which the head noun is focussed.

tituents with a non-numeral modifier but not with a numeral modifier. The reason for this difference is due to the fact that numerals are morphologically singular (cf. **kettök* 'two-pl') but are semantically specified for plural (except *egy* 'one'), whereas non-numeral modifiers can always be accompanied by a plural marker.

(*V*) Not all types of noun modifiers may participate in split constituent. For example, only adjectives, numerals, and some quantified constituents, but not demonstratives or universal quantifiers, are allowed. It appears that in split constituents only *N'-complements,* that is, sisters of the head noun, may occur. Hence, a split constituent with the nominal demonstrative pronoun *az* 'that' yields an ungrammatical result:

(6) a. Láttam *azt* a biciklit
saw-AGR1sg that-ACC the bike-ACC
'I saw that bike.'

b. *Biciklit láttam *azt*
bike-ACC saw-AGR1sg that-ACC

This explains also why an NP with *az* over which a relative clause is predicated may not be split (cf. (7a), (7c)), unlike an NP which contains its adjectival variant, i.e. the N'-complement *olyan* 'such' (cf. (7b), (7d)):

(7) a. Láttam azt a biciklit aminek piros volt a kereke
saw-AGR1sg that-ACC the bike-ACC which-DAT red was the wheel-npAGR3sg
'I saw the bike which had a red wheel.'

b. Láttam *olyan* biciklit aminek piros volt a kereke
saw-AGR1sg such bike-ACC which-DAT red was the wheel-npAGR3sg
'I saw a bike which had a red wheel.'

c. *Biciklit láttam *azt* aminek piros volt a kereke
bike-ACC saw-AGR1sg that-ACC which-DAT red was the wheel-npAGR3sg

d. Biciklit láttam *olyat* aminek piros volt a kereke
bike-ACC saw-AGR1sg such-ACC which-DAT red was the wheel-npAGR3sg
'I saw bikes such which had a red wheel.'

(*VI*) Modifiers in split constituents are nominals, more precisely *nominal predicates.* This is supported by the following two pieces of evidence.

(*i*) Modifiers in split constituents are case-marked (cf. (*II*) above). In Hungarian only members of the category *N* may bear a case-marker (cf. section 3.2.1.).

(*ii*) Some modifiers have two lexical alternants, an *attributive* and a *predicative* alternant. These alternants have a different distribution. The attributive alternant may occur only attributively, that is in a single NP. The predicative alternant may be used both attributively and predicatively. In the latter case, it heads an NP or is the predicate of a predicative sentence.

Consider, for example, the Hungarian counterparts of the modifiers *small* and *two.* The attributive alternant of the modifier *small* is *kis,* and its predicative variant is *kicsi.* The attributive alternant of the numeral modifier *two* is *két,* and the predicative variant is *kettö.*

Note that only *kicsi* and *kettö* may be the head of an NP which is modified by a determiner:

(8) a. a *kis/*kicsi*
the small
'the small one'

b. a *két/*kettő*
the two
'the two people, pieces, etc.'

Attributively, both *kis* and *kicsi* may be used, although the former is more common (cf. (9a)). In predicative sentences, however, only *kicsi* yields a grammatical result (cf. (9b)). Note now that the modifier with split constituents has exactly the same lexical shape as the predicative part of the predicative sentence (cf. (9c)):

(9) a. A *kis/kicsi* fiú
the small boy
'The small boy.'

b. A fiú **kis/kicsi*
the boy small
'The boy is small.'

c. Fiút láttam **kist/kicsit*
boy-ACC saw-AGR1sg small-ACC

Both *két* and *kettö* may be combined with a head noun, although there is a semantic divergence. Attributively kettö has a specific reading (cf. (10a)). Only *kettö,* however, may be the predicate in a predicative sentence (cf. (10b)). Again, the modifier with split constituents has the same lexical form as the modifier in a predicative sentence (cf. (10c)):

(10) a. A *két/kettö* fiú
the two boy
'The two boys/the two (specific) boys.'

b. A fiú **ket/kettö*
the boy two
'The boy is two (years old).'

c. Fiút láttam **kétet/kettőt*
boy-ACC saw-AGR1sg two-ACC

Summarizing, the fact that modifiers in split constituents are case-marked and have the same lexical shape as modifiers heading an NP or the predicative parts of a predicative sentence suggest that they are nominal predicates.

(VII) With split constituents in Hungarian *no* subject-object asymmetries turn up. In the sentences (1) and (2), we saw already that an object NP may be split. The pair in (11a) and (11b) demonstrates that a modifier may also be scrambled out of a subject, i.e. nominative, NP:

(11) a. Két ember szalad
two people run-AGR3sg
Two people are running.'

b. *Ember* szalad *kettö*
people run-AGR3sg two

Concluding this section, I would like to make the following syntactic generalization on split constituents in Hungarian:

(12) The parts of split constituents are nominal predicates and display identity of morphological features (case, number, etc.)

4.6.2. *The Semantics of Noun Modification*

In this section, I would like to discuss the semantics of the sentences in (1) and (2), here repeated for convenience as (13) and (14):

(13) a. Mari (a) két biciklit (látta)/látott
Mary (the) two bike-ACC saw-AGR3sg-def/indef
'Mary saw (the) two bikes.'

b. Mari *biciklit* látott *kettőt*
Mary bike-ACC saw-AGR3sg two-ACC
'What Mary saw two of were bikes.'

c. Mari *biciklit* látott, *kettőt*
Mary bike-ACC saw-AGR3sg two-ACC
'Mary saw only bikes and there were two of them.'

(14) a. Mari (a) nagy biciklit (látta)/látott
Mary (the) big bike-ACC saw-AGR3sg-def/indef
'Mary saw (the) big bikes/(bike).'

b. Mari *biciklit* látott *nagyot*
Mary bike-ACC saw-AGR3sg big-ACC
'Mary saw bikes such that they were big.'

c. Mari *biciklit* látott, *nagyot*
Mary bike-ACC saw-AGR3sg big-ACC
'Mary saw only bikes and they were big/big ones.'

My presentation will be rather informal. For a formal approach to the semantics of these constructions, I refer to Szabolcsi (1983c).

In the sentences ((13a), (13b)) and ((14a), (14b)) the modification is *restrictive,* whereas in (13c) and (14c) it is *non-restrictive.* The latter is indicated by a comma which corresponds in speech to a pause and a comma-intonation. Non-restrictive modification in Hungarian may be compared roughly to coordination in English as in the sentence 'Mary saw only bikes and they were big' or to the afterthought, appositional construction 'Mary saw only bikes, that is, big ones'. Before we take a closer look at the semantics of these sentences, let us first consider some different types of semantic modification.

Since Kamp (1975) the following types of semantic modification have been distinguished, among others, *intersective* and *syncategorematic* modification. I will illustrate these types through the following English pair:

(15) a. That is a big butterfly b. That butterfly is big

According to Higginbotham (1985a: 563), in (15a) the attributive modifier *big* may have only a syncategorematic reading, whereas in (15b) the predicative modifier may be used both syncategorematically and intersectively. Sentence (15a) means: 'that is a butterfly, and it is big (for a butterfly)'. The adjective is taken as grading with respect to the attribute given in the head noun. The predicative modifier in (15b), on the other hand, may have both a syncategorematic and an intersective reading. In the syncategorematic reading, it has the same meaning as (15a). However, in the intersective reading (15b) means: 'the big butterfly is a thing which is big and which is a butterfly'. Thus, when the adjective is syntactically separated from N, the semantic link may also be broken. The semantics of intersective modification can be taken as expressing *simple conjunction* (cf. Higginbotham 1985a). This implies that (15b) may count as false with respect to an object for which (15a) counts as true. Hence, from this it follows that the sentences in (15) may have different truth values.

The difference between the syncategorematic and intersective reading is illustrated even clearer by taking *stacked* adjective constructions into account. Gil (1987) notes that the following English phrases are non-synonymous:

(16) a. small powerful engine b. powerful small engine

Phrase (16a) refers to an engine that is small relative to powerful engines, whereas (16b) picks out an engine that is powerful relative to small engines. Moreover, neither of the phrases in (16) is synonymous with the phrase in (17):

(17) small and powerful engine

This phrase denotes an engine that is both small and powerful relative to engines in general.

Gil attributes the reading of stacked modifier constructions to the fact that in a hierarchical structure the sequence *A A N* may possess the structure [*A* [*A N*]]. The possibility of internal structuring enables a stacked adjective construction to be interpreted hierarchically in such a way that the outermost adjective modifies the entire [*A N*]. This yields, then, the syncategorematic readings in (16). Hierarchically, the sequence *A* and *A N* in (17) may possess the internal structure of [[*A and A*] *N*]. The adjectives are embedded under the conjunction *and.* This structuring allows (17) to be interpreted hierarchically in such a way that the entire [*A and A*] sequence modifies the N. This yields the intersective reading.

Let us turn now to a discussion of the semantics of noun modification in Hungarian.

Consider first the sentences in (14). The adjective in (14a) may have only the syncategorematic reading. So, the sentence means 'Mary saw a bike, and it was big (for a bike)'. In sentences (14b) and (14c), on the other hand, the split modifiers force the intersective reading. Hence, the meanings of (14b) and (14c) may be represented with the help of the following semantic expression: $<\forall>x$[Mary saw(x) $\rightarrow$ bike(x)] & $<\exists>x$[Mary saw(x)] & big(x). According to Szabolcsi (1983a), the universal quantifier in this constituent is provided by focussing of the head noun. This implies that the variants in (14) have different truth values.

The difference in meaning between single and split NPs is also illustrated by the Hungarian equivalents of the English stacked adjective constructions in (16) and (17). Compare:

(18) a. kis erős gép
small powerful engine
b. erős kis gép
powerful small engine
c. kicsi és erős gép
small and powerful engine

The sentences (18a) and (18b) have the same readings as their English counterparts in (16a) and (16b). In both sentences the leftmost adjective takes scope over the entire [*A N*] sequence. Hence, they display a syncategorematic reading. From this it follows that the NP in Hungarian has a hierarchical structure (cf. also chapter seven).

The phrase in (18c) displays the intersective reading, similarly as its English counterpart (17) does. It denotes an engine which is both small and powerful with respect to engines in general. Split constituents with multiple modifiers also display an intersective reading:

(19) a. Gépet láttam kicsit erőset
engine-ACC saw-AGR1sg small-ACC powerful-ACC

b. Gépet láttam erőset kicsit
engine-ACC saw-AGR1sg powerful-ACC small-ACC
'I saw engines and they were small and powerful.'

Before discussing the semantics of the sentences in (13), I will first adopt a proposal made in Verkuyl (1981) on the semantics of numerals.

Verkuyl argues that categorially numerals are adjectives, i.e. N' complements, and that their semantics may be characterized on the basis of a set-theoretical approach. For example, the numeral *Q* in (13), i.e. *két/kettő,* can be said to refer to those subsets of the power set of the denotation of the noun P that contain exactly two members. In a set expression: {<P,Q>| Card ($P \cap Q$) = 2}. This implies that this type of modifier can only have an intersective reading.

However, according to Szabolcsi, even in this triple there is a subtle semantic difference caused by the fact that the head noun is focussed in (13b) and (13c). The sentence in (13a) means that Mary saw two bikes. The sentence may be still true in case Mary saw other things like two cars, one plane and so on. In sentence (13b) Mary saw two things that were bikes. In this case the sentence is false when she saw two things not having the property bike such as two cars, two planes and so on. Of course, she may have seen one car, three planes and so on. The meaning of (13c) differs from (13a) and (13b) in that everything except bikes are barred from the universe. The comma indicates that occasionally there happened to be two bikes. Again, we conclude that the variants of noun modification may have different truth values.

Summarizing, in this section I examined the semantics of modification in Hungarian. It appeared that the triples in (13) and (14) have different truth values. They have in common that modification in all three cases expresses *conjunction* representable in a set expression. Hence, we may draw the following semantic generalization on split constituents in Hungarian:

(20) Split constituents express simple conjunction

4.6.3. *Split Constituents and the Projection Principle*

Any analysis of split constituents must solve the following two problems. First, it must avoid a violation of the Projection Principle. The mapping between LS and syntax may not be one-to-many. Second, it must account for the intuition that the variants in (13) and (14) are related semantically and syntactically. Therefore, it is justified to connect them by means of a single syntactic operation.

Such an operation is provided both by a derivational approach and by a representational approach. The former assumes the existence of an independent transformational component, or, more specifically, of the rule move-α. This means that S-structure is related to D-structure by an application of this rule. The latter, however, assumes that the rule move-α is superfluous, because the intrinsic and contextual properties of NPs at S-structure are sufficient to characterize the syntactic representation.[20]

(20) Compare Chomsky (1981), Koster (1987), and Van Riemsdijk (1982b) for further discussion of derivational versus representational grammar.

The question is whether the parts of split constituents are related by means of move-α or otherwise. In this section, I will argue that the split constituents in Hungarian provide an argument for a *representational* approach, because the parts of split constituent *cannot* be related by move-α.

Let us first discuss the derivational analysis of the triples in (13) and (14), here repeated as (21) and (22):

(21) a. Mari (a) két biciklit (látta)/látott
Mary (the) two bike-ACC saw-AGR3sg-def/indef
'Mary saw (the) two bikes.'
b. Mari *biciklit* látott *kettőt*
Mary bike-ACC saw-AGR3sg two-ACC
'What Mary saw two of were bikes.'
c. Mari *biciklit* látott, *kettőt*
Mary bike-ACC saw-AGR3sg two-ACC
'Mary saw only bikes and there were two of them.'

(22) a. Mari (a) nagy biciklit (látta)/látott
Mary (the) big bike-ACC saw-AGR3sg-def/indef
'Mary saw (the) big bikes/(bike).'
b. Mari *biciklit* látott *nagyot*
Mary bike-ACC saw-AGR3sg big-ACC
'Mary saw bikes such that they were big.'
c. Mari *biciklit* látott, *nagyot*
Mary bike-ACC saw-AGR3sg big-ACC
'Mary saw only bikes and they were big/big ones.'

Horvath (1986: 29; 83) proposes a derivational analysis of split constituents. In Horvath's acccount the (a)-sentences in (21), and (22) are taken as the underlying structures for their counterparts in (b) and (c). These sentences are derived by applying Quantifier Float and Topicalization respectively. Move-α scrambles the modifier out of its base-generated position and leaves a trace in the modifier position of the NP. This analysis of split constituents does not violate the Projection Principle. However, I will discuss the following *morphological* (cf. *I*), *syntactic* (cf. *II-III*), and *semantic* (cf. *IV*) anomalies arising with this type of derivation.

(*I*) The derivational analysis leaves some *morphological* dichotomies unexplained between the split and unsplit variants. First, it is unclear where the case-marker on the modifier in the split variant comes from. Second, this problem appears also with the plural marker on the head noun in (4d). Recall that in the underlying structure only singular head nouns are allowed when the head noun is in construction with a modifier expressing quantity.

Third, the derivational analysis must allow for the formation of new lexical predicative stems after scrambling the attributive modifier out of its NP, for example, *két/kettö,* and *kis/kicsi* in (9), and (10). If the triples in (21) and (22) are indeed related by an application of move-α, then this contradicts the *Lexical Integrity Hypothesis* (cf. Lieber (1980)) which states that NPs are base-generated in their fully inflected forms.

(*II*) Horvath (1985, section 1.3.) refers to split constituents as 'Quantifier Float'. This term suggests, however, that a generalization is missed. Not only numerals or

quantifiers but also adjectives may appear in the split variant (cf. (22)). The question is why only these modifiers may be scrambled out of their NP.

Horvath further claims that Quantifier Float obeys an adjacency requirement. She cites the following examples to illustrate this:

(23) a. Mari nem mutatta be az új diákot mindegyik tanárnak
Mary not showed-AGR3sg in the new student-ACC each teacher-DAT
(Horvath 1985: 27, (19a))

b. Mari nem mutatta be az új diákot a *tanároknak mindegyiknek*
Mary not showed-AGR3sg in the new student-ACC the teachers-DAT each-DAT
(Horvath 1985: 27, (19b))

According to Horvath (1985: 27), the QP *mindegyik* 'each' occurs either in the specifier position of NPs (cf. (23a)), that is, on a left branch within NPs, or outside the NP as a result of Quantifier Float (cf. (23b)).

Horvath lists the following properties of Quantifier Float including (*i*) the quantifier exhibits case-marking identical to the head noun (p. 27), (*ii*) the head noun must be plural (p. 27), (*iii*) the quantifier must be adjacent to the NP it modifies (p.28), (*iv*) the QP must occur to the right of its NP (p. 82, fn. 15), (*v*) absence of subject-object asymmetries (p. 30), and (*vi*) the Quantifier Float also has a right dislocated variant with the QP base-generated in the right dislocated position. Such structures are ungrammatical in case the right dislocated QP is in the scope of a negation operator (NEG) (p. 82, fn. 15).

Horvath argues that her SVO-hypothesis of the Hungarian in combination with the properties of Quantifier Float listed above can account for the difference between the following two structures:

(24) a. ?*...NEG V NP-DAT... QP-DAT
Horvath (1986: 28, (21a))

b. ...NEG NP_i-DAT V... t_i QP-DAT
Horvath (1986: 28, (21b))

In an SVO-structure non-subject NPs are base-generated postverbally. According to Horvath, the reason why (24a) is ungrammatical and (24b) is not involves a violation of the adjacency requirement on Quantifier Float in the former. The latter escapes the violation of this requirement since the head noun has been subject to move-α and is (via its trace) adjacent to the QP. Horvath claims thus that the floated QP must be right-adjacent to the head noun or its trace.

In Szabolcsi (1983c), however, a number of examples are presented which are not in accordance with this claim. Of course, they could fall under Horvath's transformational approach. The crucial example in favour of Horvath's adjacency requirement is provided by properties of structures as (24a). Let us carefully examine this case.

Horvath observes that (24a) is not an instance of a right dislocated structure. Hence, it cannot be ruled out by her rule that right dislocated QPs may not be in the scope of NEG (cf. Horvath (1985), 82, fn.15)). Therefore, she concludes that the reason for its ungrammaticality must be a violation of the adjacency requirement. Szabolcsi (1983c, fn.8), however, observes that sentences with a non-dislocated QP in the scope of a NEG are perfectly grammatical:

(25) Biciklit nem látott Mari *kettőt*
bike-ACC not saw-AGR3sg Mary two-ACC
'What Mary didn't see two of were bikes.'

In this sentence the head noun and the floated QP are not adjacent. This casts doubt on Horvath's claim that an adjacency requirement is operative with Quantifier Float.

A further question which Horvath does not discuss is why only bare Ns undergo Quantifier Float.

(*III*) It is a well-known fact that languages possessing floating quantifiers display subject-object asymmetries with this phenomenon. (See, for example, Haig 1980 for such asymmetries in Japanese). With split constituents in Hungarian, however, no subject-object asymmetries arise (cf. section 4.6.1.(*VII*)).

(*IV*) It remains unclear under a derivational analysis why the split variant may only have an intersective reading. Because of the trace in the modifier position, the NP-configuration remains unaffected. Hence, the syncategorematic reading should be available in case of a split NP as well.

From the problems listed in (*I*)-(*IV*), I conclude that a derivational analysis of split constituents makes the wrong predictions and leaves open a number of questions. An alternative analysis of this type of constituents is provided by a representational approach to grammar. Below I will elaborate such an analysis of split constituents along the lines of Higginbotham's (1985a; 1986) theory of θ-*discharge.*[21] Before doing so, let us first consider the concepts relevant for our approach.

Higginbotham (1985a) proposes the following redefinition of the θ-*criterion* (cf. 3.2.(2)):

(26) θ-*Criterion:* a. Every argument is assigned one and only one θ-role
b. Every θ-position is discharged (uniquely)

The original second part of the θ-criterion (cf. 3.2.(2b)) is now replaced by (26b) which is more general. The elimination of open θ-positions in the θ-grid of lexical items is not only restricted to arguments under this approach.

Higgingbotham distinguishes the following types of θ-discharge:

(27) a. θ-*marking,* exemplified by pairs consisting of a predicate and one of its arguments
b. θ-*identification*, exemplified in simple adjectival modification as in *white wall* interpreted as 'white(x) and wall(x)'
c. *Autonymous* θ-*marking,* where the value assigned to the open position in the θ-marker is the attribute given by its sister
d. θ-*binding*, exemplified by determiners or measure-words and their nominals, as in *every dog*, interpreted as 'for every x such that dog(x)'

These modes of discharge are the primitive semantic operations of structural meaning which are all controlled by the configuration of *government* (mostly identifiable with sisterhood). θ-marking covers the nonmodificatory, or simple case of θ-discharge. The others types refer to a modification relation.

Consider an example of each of the latter type. Let us first discuss θ-*identification.*

Bare nominals are open constructions. This is supported by the fact that nominals can serve as predicates in many languages. Adjectives must have an open position

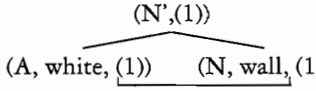
(21) Pica (1987) applies this theory to reflexive anaphors. According to Pica, the fact that reflexive anaphors must be bound by an antecedent is due to the property that they have an open position in their syntactic representation which must be saturated.

as well since they may also function as a predicate. Hence, we may assign nominals and adjectives the following θ-grids as part of their lexical entries ((1) indicates that there is an undischarged role associated with the predicate):

(28) a. *nominal,* [-V, +N], (1) b. *adjective,* [+V, +N], (1)

The semantics of the phrase *white wall* is expressed by a simple conjunction: a white wall is a thing that is white and a wall. In this phrase, some position in the adjective is identified with the nominal position. The θ-structure of *white wall* may be represented in the following diagram:

(29)



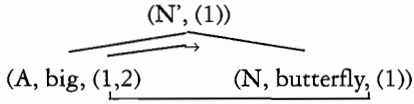
The open position of the adjective is discharged under this identification, indicated by the connecting line. We can compare its structure to that of building up a compound *Fx* & *Gx* and then identifying *x* and *y*.

Let us consider now an instance of *autonymous* θ-*marking*. Consider again (15a), here repeated as (30):

(30) That is a big butterfly

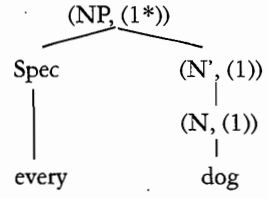
This phrase can be paraphrased as follows: that is a butterfly, and it is big (for a butterfly). In this paraphrase, the head noun is an argument of the adjective. So, this category serves to discharge two θ-positions in a syncategorematic adjective-noun construction. One by identification and the other by θ-marking of the noun itself by the adjective. This latter mode is called autonymous θ-marking, indicated by an arrow in diagram (31). The tail of the arrow is at the position of the θ-marker and its head abuts the point marked:

(31)



Higginbotham notes that head nouns do not take arguments when they form NPs. What happens instead is that the position (1) in (28a) is accessible to Spec, which acts as a binder. There must be some binder, and there can not be two. This mode of θ-discharge is referred to as θ-*binding*. The θ-*structure* of, for example, *every dog* might be depicted as follows (the asterisks indicates that the open position in N' is not open in NP):

(32)



Having discussed several modes of θ-discharge, let us turn now to a representational analysis of split constituents.

I will first examine the representations (21a) and (22a). Recall that numeral modifiers always display an intersective reading (cf. section 4.6.2.). So the modification in (21a) is an instance of θ-identification. We observed that (22a) is a case of syncategorematic modification. Analogously to (30a), we may handle this sentence by the combination of θ-identification and autonymous θ-marking. Let us turn to (21b) and (22b).

Suppose we assign for example the head noun *biciklit* the status of direct object argument in these sentences. Either it is in a complement position itself, or it is related to this position by scrambling. This has two consequences.

First, the Projection Principle is satisfied, because the transitive verb *lát* has now two arguments, a subject and an object. Second, the modifiers *kettőt* and *nagyot* turn into adjuncts. Accordingly, I assume that they are base-generated in a non-A-position, as any other adjunct is. This accounts then for the fact that the parts of split constituents display freedom of word order. The question arises then how the parts of split constituents are related under a representational approach.

The semantics of split constituents is characterized by simple conjunction (cf. (20)). From this it follows that the parts of these constituents must be related by means of θ-identification. The question to answer is how this relation is set up and how it is restricted.

Nominals and adjectives have an open position in their θ-grid which must be discharged. This covers the fact why only certain types of modifiers (adjectives, numerals, some quantifiers) may participate in split constituent constructions, namely, exactly those which may function as predicate nominals, and thus may be open structures.

The fact that the parts of split constituents have this property also provides an explanation for the observation that they must be bare. Modification by a definite or indefinite determiner would close the structure, i.e. eliminate its θ-role from the grid, by the mode of θ-binding. Hence, θ-identification would be blocked as a violation of the θ-criterion (cf. (26b)).

Szabolcsi (1986b: 48) notes an interesting apparent exception to this restriction. Szabolcsi observes that if the separated modifier is in the superlative it may be modified by a determiner:

(33) Zöld lóval itt találkoztam a legszebbel
green horse-INSTR here met-AGR1sg the prettiest-INSTR
'I met a prettier green horse here than anywhere else.'
*'As for green horses, it was here that I met the prettiest of them, i.e. the prettiest green horse that there is.'

Observe from the glosses that a superlative adjective modified by a definite article may only be separated from the head noun in the comparative reading. This sentence is ungrammatical in the absolute reading of the superlative.

Szabolcsi claims that the superlative phrase in the absolute reading is in the Spec of the NP, whereas the superlative phrase is NP-internal in the case of the comparative reading. Her conclusion fits in nicely with our result. In the absolute reading the NP would be closed by a binder in the Spec of the NP. In the comparative read-

ing, although there is a determiner present, the NP counts as an open structure with an undischarged θ-position. Hence, the superlative phrase may be available for split constituents only on the comparative reading.

Higginbotham assumes that θ-identification is restricted by government. One part of split constituents must be base-generated in a non-A-position outside a maximal projection by assumption. Therefore, θ-identification in these cases cannot be restricted by government. Instead I will assume that this type of θ-discharge between the parts of split constituents is conditioned by a weaker structural condition than government, namely, by *c-command.* C-command is the minimal structural condition two mutually dependent constituents generally have to obey. It is always respected in split constituents because one of the parts is in a non-A-position from where it can c-command the part in a complement position. This covers then the fact that no subject-object asymmetries turn up with split constituents.

θ-identification is further restricted by a morphological licensing condition, i.e. *identity of morphological features* (case, number, etc.). This depends on the different functions morphological markers may have in a language. In Hungarian, case-markers may act as an attribute relater in split constituents. A case-marker indicates that an adjunct is predicated of the head noun. The optional agreement of plural features does not form an obstacle under this analysis (cf. 4.6.1.(IV)).

If this analysis of noun modification in Hungarian is on the right track, we expect that the following predictions about the possibility of "splitting" single NPs will be borne out. In case the semantics of a modifier-noun combination cannot be captured by a simple conjunction, a split constituent is not allowed, or to put it differently, if it is not possible to take the combination of adjective and noun as having as its denotation the intersection of the set denoted by the adjective with the set denoted by the noun. This arises at least in the following two cases.

(*i*) Modification is interpreted as a combination of θ-identification and autonymous θ-marking. Consider the following pair:

(34) a. Mari nagy bolhát látott
Mary big flea-ACC saw-AGR3sg
'Mary saw a big flea.'

b. *Mari bolhát látott nagyot
Mary flea-ACC saw-AGR3sg big-ACC

Sentence (34a) means: 'Mary saw a thing that was a flea and it was big for a flea'. This reading is a case of a combination of θ-identification with autonymous θ-marking. In sentence (34b), the modifier is separated from its head noun. According to generalization (20), the semantics of such constructions is captured by simple conjunction. Therefore, the sentence should mean: 'Mary saw a thing that was a flea and it was big'. However, this is not in coherence with the properties of fleas. Even big fleas are not big creatures. Hence, the ungrammaticality of (34b).

(*ii*) Modification is interpreted as a case of autonymous θ-marking only. Examples of the latter are phrases like *former president* and *alleged murder.* Adjectives such as *former* and *alleged* cannot have as denotation the intersection of any such sets. For example, *former president* cannot be analysed as the intersection of the set of presidents with the set of things *that are former*. It is easy to see that this latter phrase is meaningless. Compare the following sentences:

(35) a. Mari látta az előző elnököt
Mary saw-AGR3sg the former president-ACC
'Mary saw the former president'
b. *Mari elnököt látott előzőt
Mary president-ACC saw-AGR3sg former-ACC

In sentence (35a) the phrase 'former president' designates a person whose presidency is former. This is not a case of intersective modification (cf. Higginbotham 1985a: 567, who suggests to analyse this case by adopting a temporal positions in the θ-grids of nouns). Hence, as (35b) shows, it is impossible to have the split variant.

Consider the following pair:

(36) a. Mari látta az állitólagos gyilkost
Mary saw-AGR3sg the alleged murder-ACC
'Mary saw the alleged murder.'
b. *Mari gyilkost látott állitólagosat
Mary murder-ACC saw-AGR3sg alleged-ACC

Sentence (36a) is a case of autonymous θ-marking (cf. Higginbotham 1985a). An alleged murder is true of things alleged to be a murder. Again, it is not possible to form a split constituent construction (cf. (36b)), as this presupposes an intersective reading of the modifier-noun combination.

Summarizing, I have discussed the syntax and semantics of noun-modification in Hungarian. I have isolated the syntactic and semantic conditions under which one of the variants of noun-modification, the split constituent, appears (cf. (12) and (20)). Further, I have considered two analyses which relate the split constituent to the unmarked single constituent, a derivational and a representational one. Neither of them violates the Projection Principle, because the biuniqueness requirement on this principle is obeyed.

I have argued that a representational analysis makes the better predictions. The reason for this is that the derivational approach assumes that the parts of split constituents are related by trace-binding. At the position of the modifier a trace is postulated. This assumption causes morphological, syntactic, and semantic anomalies.

The representational approach, on the other hand, assumes that some positions in the grids of the parts of split constituents are related. This type of binding is distinct from trace-binding. A position in a grid is not a formative of the grammar but rather a part of the lexical entry of a predicator whose grid it is. Therefore, the above anomalies are avoided under a representational analysis.

Let us turn now to a discussion of split constituents in other languages.

4.6.4. *Split Constituents and the Theory of Grammar*

In the literature, split constituents have been discussed most extensively for Warlpiri and German. As we will demonstrate below, the properties of split constituents in these languages coincide largely with the ones of Hungarian. Therefore, it is suspicious that they give rise to widely different theoretical views. First, I will discuss split constituents in Warlpiri and then I will turn to this phenomenon in German.

(I) Split Constituents in Warlpiri

Hale (1983) argues that the Projection Principle applies in *Warlpiri* only at LS. Under this assumption, a many-to-one linking from PS onto LS is allowed. Consequently, split constituents may appear unrestrictedly.

The following sentences exemplify some instances of this phenomenon in Warlpiri:

(37) a. *Wawirri* kapirna pantirni *yalumpu*
kangaroo Aux spear-nonpast that
'I will spear that kangaroo.'
(Hale 1983: (4))

b. *Malikirli* Φ- ji yarlkurnu *wiringki*
dog-ERG perf 1obj bite-past big-ERG
'The/a big dog bit me.'
'The/a dog me and it was big.'
(Hale 1983: (39))

Word order is free in these sentences, apart from Aux, which is usually in second position in Warlpiri.

The composing parts of an NP in English may appear linearly non-adjacent in a Warlpirian clause. For example, in (37a) the restrictive determiner *yalumpu* modifies the noun *wawirri* as in the English translation. According to Hale, this discontinuous pair forms an expression corresponding to that represented by the single syntactic constituent *wawirri yalumpu* in (38):

(38) *Wawirri yalumpu* kapirna pantirni
kangaroo that Aux spear-nonpast
'I will spear the kangaroo.'
(Hale 1983: (5))

Let us consider now whether split constituents in Warlpiri obey the same restrictions as the ones in Hungarian.

I formulated the syntactic restriction (12) on the occurrence of this phenomenon, here repeated for convenience as (39):

(39) The parts of split constituents are nominal predicates and display identity of morphological features (case, number, etc.)

Nash (1980) and Hale (1981) observe that the parts of split constituents in Warlpiri must have the same categorial and morphological features (N, case, number) as well. For example, in (37a) both parts are in the absolutive singular, and in (37b) they are marked ergative singular. Hence, split constituents in both Warlpiri and Hungarian display *identity of morphological features.*

Several authors (cf. Nash 1980, Hale 1981; 1983, and Simpson 1983) claim that the category N includes both nominals and adjectives in Warlpiri. There are no formal morphological and syntactic properties which distinguish these parts of speech. Nash (1980: 15), for example, points out that adjectives are in fact nominals that prefer a reading which has an argument position in it. This is illustrated by the following sentence:

(40) Pakarni kapala maliki *witajarrarlu*
strike-nonpast Aux-pres-3dual dog small-dual-ERG
'The two small ones (children, say) are striking the dog.'
(Hale 1981: (31))

Note that the adjective *wita* 'small' may receive the interpretation of a full NP in this sentence.

We may conclude then that both adjectives and nominals in Warlpiri may function as nominal predicates. In fact, any part of an NP in English may be turned into an independent NP in this language. From (37a), it is clear that even a determiner such as *that* displays this property, since it participates in split constituents.

Hence, there seems to be a correlation between the ability to promote a modifier into a predicate and the participation of that modifier in split constituents. Languages may differ with respect to this ability. For example, modern English does not display split constituents of the type discussed here. Nor may modifiers head an NP. As a consequence, in an elliptical NP the missing head must be represented by one:

(41) a big *(one)

In Hungarian and Warlpiri, modifiers may be promoted into predicates, although the group of modifiers participating is more restrictive in Hungarian. Hungarian allows these modifiers to be only adjectives, numerals, and some quantifiers, whereas Warlpiri allows all modifiers to become nominal predicates.

This ability to promote modifiers into predicates might then be a parametric difference among languages. English and Warlpiri are on the ends of the scale, while Hungarian is somewhere in the middle. English has no split constituents, in Warlpiri this phenomenon appears freely, and in Hungarian split constituents do occur but not as freely as in Warlpiri. In sum, there is no difference between Hungarian and Warlpiri in the syntactic status of the split parts. In both languages, they are *nominal predicates* which head an NP.

Let us discuss now whether the semantics of split constituents in Warlpiri coincides with the semantics of these constituents in Hungarian. Recall that (12), here repeated as (42), captures the semantics of Hungarian split constituents:

(42) Split constituents express simple conjunction

According to Hale (1983), split constituents in Warlpiri may receive at least two interpretations. Consider again sentence (37b), here repeated as (43):

(43) *Malikirli* Φ- ji yarlkurnu *wiringki*
dog-ERG perf 1obj bite-past big-ERG
'The/a big dog bit me.'
'The/a dog me and it was big.'

On one reading of this sentence, the expression *wiringki* is taken as a modifier of *malikirli,* constituting an expression which corresponds to the single constituent *maliki wiringki* in the following sentence:

(44) *Maliki wiringki* Φ- ji yalkurnu
dog big-ERG perf 1obj bite-past
'The/a dog bit me.'

Hale refers to this reading as the 'merged' interpretation. That the subject in (44) is a single constituent is shown not only by the position of Aux but also by the manner in which the case-category of this expression is marked. The ergative suffix appears on the final subconstituent only. On the other reading of (43), *wiringki* is simply predicated of *malikirli.* It receives an unmerged interpretation.

The parallelism between split constituents in Hungarian and Warlpiri breaks down at this point. Hungarian split constituents have only what Hale calls an unmerged interpretation (cf. (42)). There is, however, some reason to be careful with the claim that one of the interpretations of the split constituent in (43) is synonymous with the interpretation of the single expression in (44). Hale himself (1983: fn.2) notes that the role of word order in interpretation is an aspect of Warlpiri which is still very much in need of investigation. Furthermore, McGregor (1989), in a paper on split constituents in Gooniyandi (another aboriginal Australian language related to Warlpiri) emphasizes that single and split constituents have different semantic, pragmatic, and phonetic properties.

Summarizing, split constituents in Hungarian and Warlpiri display the same syntactic properties. The parts of split constituents may be linked only under identity of morphological features, and they are nominal predicates. Semantically, Hungarian and Warlpiri seem to diverge with this phenomenon. Hungarian split constituents do not exhibit a merged interpretation.

In my view, these syntactic parallelisms between split constituents in Hungarian an Warlpiri do not justify a radically different analysis. If these constituents do not violate the Projection Principle in Hungarian, then the null-hypothesis should be that they do not in Warlpiri either.

An analysis of Warlpiri which respects the Projection Principle has been elaborated in Jelinek (1983). Jelinek argues that the clitic pronouns in Aux serve as verbal arguments which satisfy the Projection Principle. As a consequence, nominals are not verbal arguments but are adjuncts coindexed with these arguments. There is nothing which prohibits the binding of the parts of split constituents along the lines of section 4.6.3. The precise elaboration of this, however, is beyond the scope of this study.

(II) Split Constituents in German

Below, I will discuss split constituents in *German.* In my discussion, I will heavily rely on observations made in Bayer (1987), Fanselow (1987b), and Van Riemsdijk (1987). Split constituents in German have a number of properties in common with such constituents in Hungarian. Hence, it is attractive from a theoretical point of view to analyse them in a similar way.

Van Riemsdijk (1987) argues for a derivational approach. In order to do so, Van Riemsdijk proposes to extend derivational grammar with the theory of regeneration. I will argue, however, that Van Riemsdijk's analysis is rather defective in that it makes a number of *ad hoc* claims and incorrect predictions. Before entering this theoretical debate, let us first examine some of the relevant properties of split constituents in German.

(i) According to Van Riemsdijk, this phenomenon in German is formed by topicalizing the head noun which is an N' and leaving behind the determiner in the sour-

ce position. Therefore, Van Riemsdijk refers to this construction type as *Split Topicalization* (ST). Compare the following example:

(45) *Bücher* habe ich *keine* mehr
books have I none more
'As for books, I don't have any more.'
(Van Riemsdijk 1987: (1))

The head noun *bücher* is in the preverbal topic position which is identified as the Spec of CP position, and the stranded determiner is in the object position in this sentence. Recall split constituents in Hungarian are not so positionally restricted as ST in German (cf. 4.6.1.(I)).

Van Riemsdijk observes further that the topic NP must be indefinite. It may not be modified by an overt definite or indefinite article, except by *ein* in some southern varieties of German (cf. *(iv*A) below). The split source must be indefinite but is not otherwise restricted. Split constituents in Hungarian are restricted by a *definiteness effect* as well (cf. 4.6.1.(III)).

(*ii*) ST requires identity of number and case agreement. In the following sentence, both parts must be in the accusative singular:

(46) *Einen/*ein* *Wagen* has er sich noch *keinen* gekauft
a-ACC/NOM car has he refl yet none bought
(Van Riemsdijk 1987: (20))

Recall that split constituents in Hungarian display *identity of morphological features* as well, at least with reference to the case-marker (cf. section 4.6.1.(II)).

(*iii*) Some determiners, such as *kein* 'no', bear different *adjectival inflection* depending on whether they are in an independent elliptical NP or whether they are in construction with a head noun. When they act as an independent NP their inflection switches from weak (cf. (47b)) to strong (cf. (47a)):

(47) a. Er hat *keines*/*kein
he has no

b. Er hat *kein*/*keines Geld
he has no money
(Fanselow 1987b: (43))

Such a determiner in the source position of ST must take the inflection of the independent form, that is, it must appear with strong inflection:

(48) *Geld* hat er *keines*/*kein
money has he no

Fanselow (1987b) argues that in case the modifier appears in an elliptical NP, it has not become a noun. The reason for this is that even in such an NP the modifier retains its adjectival properties.

There are three-classes of case-number-gender endings for adjectives, the so-called 'strong', 'weak', and 'mixed' systems of inflection. Choice among them is triggered by the respective determiner:

(49) a. Ein *roter* Apfel
a red apple

b. Der *rote* Apfel
the red apple

Nouns, on the other hand, have just one class of case-number-inflection:

(50) a. Ein Bote
a herald
b. Der Bote
the herald
(Fanselow 1987b: (46))

Note now that in NPs without overt nominal heads, adjectives retain the three-valued system of inflection:

(51) a. Ein interessanter
an interesting
'An interesting one'
b. Der interssante
the interesting
'The interesting one'
(Fanselow 1987: (47))

(*iv*) The parts of ST cannot appear as single NPs in some cases. This implies that there is in these cases *no* source to which the parts of ST can be related under a movement analysis.

(*A*) Van Riemsdijk observes that some varieties of southern German allow the head noun in topic position to be modified by an independent determiner. Van Riemsdijk refers to this as *determiner overlap*. With this phenomenon, the determiner in the topic part is always the simple indefinite article, that is, *ein* for singular count nouns and zero for singular mass nouns and plurals. Compare, for example:

(52) *Einen Wagen* hat er sich noch *keinen* leisten können
a car has he refl yet none afford could
'As for cars, he has not been able to afford one yet.'
(Van Riemsdijk 1987: (4))

In this sentence, both the topic and the source position contain a determiner. The split constituent cannot be derived from an underlying source which contains both determiners, because such a single NP does not exist:

(53) a. *Einen keinen Wagen b. *Keinen einen Wagen

(*B*) Another case in which the parts of ST cannot be related to a single NP is with the word *welch-*. As an independent NP, it has the meaning of an existential quantifier (cf. (54a)). When it is part of a single NP, however, *welch-* cannot have this meaning (cf. (54b)):

(54) a. *Geld* frage ich mich ob er *welches* hat
money ask I me whether he some has
'I wonder if he has some money.'
b. *Ich frage mich ob er *welches Geld* hat
'I wonder if he has some money.'
(Fanselow 1987b: (40))

(*v*). Van Riemsdijk (1987: 6) and Bayer (1987) argue that the parts of ST are *complete* NPs.

First, determiner overlap: both the topic and source part are turned by the determiner into a full NP (cf. (52)).

Second, the determiner has strong inflection in ST just as in independent elliptical NPs (cf. (51)). Recall that the parts of split constituents in Hungarian display

this property too. They have the same lexical forms as independent elliptical NPs or the predicative parts of predicative sentences (cf. section 4.6.1.(VI))

Third, nouns which cannot function as independent NPs may not participate in ST. Hence, the singular unmodified noun *Foto,* unlike its plural counterpart, yields an ungrammmatical result with ST:

(55) a. Fotos sehe ich viele
photos see I many
'I am seeing many photos.'
b. Ich sehe Fotos
I see photos
'I am seeing photos.'
c. *Foto sehe ich das
photo see I that
d. *Ich sehe Foto
I see photo
(Fanselow 1987b: (45))

(*vi*) Joseph Bayer (personal communication) informs me that ST is subject to a similar semantic restriction as split constituents in Hungarian (cf. (20)). A non-intersective modifier may not participate in ST. Therefore, (56a) with the intersective modifier *rothaarig* is grammatical, unlike (56b) with the non-intersective modifier *angeblich:*

(56) a. *Mörder* hat er *einen rotharigen* getroffen
murder has he a redhaired met
'As for a murder, he has met a redhaired one.'
b. **Mörder* hat er *einen angeblichen* getroffen
murder has he a alleged met

(*vii*) According to Van Riemsdijk, the meaning of the word *welch-* is dependent on its syntactic context. When it is part of an NP modifying a head noun, it has the meaning of *which* (cf. (57a)), but it has the meaning of an existential quantifier when it is elliptical (cf. (57b)). (If the existential reading is preserved in a single NP *welch-* must be pefixed with *irgend-*). Van Riemsdijk observes now that *welch-* may only have the existential reading *when* it participates in ST (cf. (57c)):

(57) a. *Welche* unbeschädigten Exemplare hast du?
which undamaged copies have you
(*welche* = which)
b. Hast du *welche?*
have you any
(*welche* = some)
c. Unbeschädigte Exemplare habe ich kaum noch *welche*
undamaged copies have I hardly still any
(*welche* = some)

(*viii*) According to Van Riemsdijk, ST obeys the diagnostics of move-α, since it is sensitive to island constraints. It may not violate the Wh-island Constraint, the Complex Noun Phrase Constraint, and it does not allow Preposition Stranding. Further, it displays ECP-effects with extraction from the left-branch and it displays reconstruction effects.

Van Riemsdijk observes a paradox now. Several properties of ST such as the fact that there is no underlying source for a movement analysis in the case of determiner overlap (cf. (*ivA*)) and with the word *welch-* (cf. (*ivB*)), and the fact that both parts of ST are complete NPs (cf. (*v*)) suggest that ST cannot be derived by move-α. On the other hand, it obeys the diagnostics of movement. In order to escape this paradox,

Van Riemsdijks invents the theory of regeneration which filters the application of move-α.

Let us summarize the essence of this theory. Van Riemsdijk allows move α to affect any category on the X'-projection. According to Van Riemsdijk, ST involves an instance of N'-movement which leaves a trace in the source position.

Van Riemsdijk formulates an S-structure filter which does not allow S-structure representations containing an X'-category which is not dominated by its maximal projection. This filter permits regeneration of the X'- projection in topic position into a full-fledged NP.

Regeneration may be followed by the partial relexicalization of the regenerated structures. The relevant morphosyntactic features such as [count], [gender], [number], and [case] which are for the most part inherent features of the head noun, are used to determine the lexical form of the determiner. A recoverability requirement on relexicalization accounts for determiner overlap.

Regeneration and relexicalization are subject to parametric variation, since some languages such as modern English do not allow split constituents, and some dialects of German do not allow determiner overlap with ST.

Van Riemsdijk states that this derivational theory of ST is both theoretically and empirically superior to a representational account. However, regeneration runs into the following anomalies.

(*I*) It is not obvious why the machinery of regeneration applies at all in case of ST. Van Riemsdijk assumes that ST is an instance of N'-movement. Some of its properties, however, suggest that both parts are full NPs (cf. (*v*)). This is also acknowledged by Van Riemsdijk (1987: 6) himself. In other words, it remains unclear why the source NP should contain an N'-gap.

Alternatively, it could be assumed that the determiner/modifier in source position heads the remnant NP, such as with split constituents in Hungarian (cf. section 4.6.3.), or it could be assumed that the head of the source NP is small *pro* (cf. Fanselow 1987b).

According to Fanselow, the latter alternative also explains the switch of weak to strong inflection in elliptical independent NPs and the source NP in ST. Only the strong adjectival inflection can license *pro*. Therefore, in languages without strong adjectival inflection, like English, *pro* has to be spelled out in elliptical NPs:

(58) A in german [illegible]

Van Riemsdijk claims that under the movement theory the contrast in (60) can immediately be reduced to the principles that account for the ordering restrictions on the adjectives in the non-split NP in (59). Although the correlation between these pairs may be accounted for by regeneration, it is also in agreement with an alternative theory along the lines of section 4.6.3. Since the predication relation between the parts of ST qualifies full NPs.

But let us turn now to ordering restrictions with NPs in Hungarian. Compare the Hungarian counterparts of (59):

(61) a. egy új *amerikai* autó
a new American car
b. *egy *amerikai új* autó

However, contrary to German in Hungarian the internal order of modifier within a single NP does not have to be preserved with split constituents:

(62) a. Amerikai autót vettem *ujat*
American car-ACC bought-AGR1sg new-ACC
'I bought an American car such that it was new.'
b. *Új* autót vettem *amerikait*
new car-ACC bought-AGR1sg American-ACC
'I bought a new car such that it was american.'

Thus, the correlation in the German pairs (58) and (59) does not turn up in these Hungarian pairs. If the correlation in German is an argument in favour of move-α in ST, then the absence of such a correlation in Hungarian is an argument against this rule in Hungarian split constituents.

(*III*) Regeneration runs into an ordering conflict with lexical insertion. Van Riemsdijk (1987: fn.5) assumes that relexicalization applies at or after S-structure, because of the morphological form of determiners which participate in the inflection switch such as *kein* (cf. (*iii*)). However, Van Riemsdijk (p.29) also assumes that lexical insertion must apply at D-structure, since otherwise the relative order of adjectives (cf. (*II*)) cannot be determined. Such principles have to refer to the lexical content of adjectives. Thus, lexical insertion must take place at D-structure to account for the relative order of adjectives, but it may not apply at D-structure otherwise the morphological form of some determiners cannot be predicted. To assume, however,

ite article is the unmarked form of the nominal determiner which must be overtly represented in some dialects. It seems to me that this spelling out of the indefinite article does not prevent the topic NP to participate in the unification of the parts of ST. Hence, this fact remains neutral with respect to the choice between movement and base-generation.

Summarizing, ST in German has a number properties in common with split constituents in Hungarian. Some of these properties conflict with a movement analysis. Van Riemsdijk (1987), however, extends the derivational theory with the theory of regeneration in order to account for ST. I have pointed out that this theory makes some wrong predictions, both in the case of German ST and Hungarian split constituents. Therefore, it should be treated with some scepticism. However, the elaboration of a representational analysis of ST lies beyond the scope of this study.

4.7. Conclusions

In this chapter, I presented empirical evidence from Hungarian for the hypothesis that the Projection Principle holds in the mapping from LS onto syntax. Therefore, it is not justified to parametrize the Projection Principle in order to derive some of its apparent 'non-configurational' properties, such as relative free word order or split constituents. This chapter supports the claim that the Projection Principle is a universal principle.

The Projection Principle seems to be violated by omitted pronouns (cf. section 4.2. and 4.5.) and by split constituents (cf. section 4.6.). In the former case, the correspondence between LS and syntax is one-to-null, and in the latter case this correspondence is one-to-many.

However, I argued that the position of omitted pronouns is taken by small *pro*. The presence of this empty category in Hungarian follows from the fact that it displays the same distribution as its overt counterpart, and that there is a functional split between *pro* and its overt counter-part with some syntactic phenomena.

Split constituents may appear only under highly specific syntactic and semantic restrictions. This implies that they are are rather marked. In fact, they are "saved" by an interaction of θ-theory with Case theory, more precisely, with the properties of overt case-markers in Hungarian. So, these phenomena do not question the hypothe-

Left-dislocated NPs and embedded clauses cannot receive Case- and θ-features directly from the governing verb. Therefore, I assumed that personal and demonstrative pronouns have the ability to transfer these features to (clausal) antecedents.

Split constituents may be derived under a derivational or a representational approach to grammar. I argued that a representational approach makes the better predictions with this phenomenon in Hungarian.

The phenomena discussed in this chapter do not only support the hypothesis that the Projection Principle applies between LS and PS but they may also give us some insight into the way lexical information is projected. For example, θ-governed lexical case must be visible at surface structure. This appeared from personal pronouns with case-stems, double-faced lexical items, the demonstrative/personal pronoun-switch in the formation of embedded clauses, and left-dislocated structures. What seems to be projected onto syntax with these phenomena is Case. The Projection Principle is category blind in these cases. The properties of transfer systems allow then the different types of categorial constituents in syntax.

If this is correct, it provides an argument for the autonomy of LS. It would be worth investigating whether LS is an independent module. A more complete elaboration of such puzzles will have to await, however, further research.

In this chapter, I have argued that the relation between LS and syntax is subject to a biuniqueness condition. I have, however, not argued for the particular formulation of the Projection Principle in 4.1.(2):

(1) The LS must be represented categorially at each level of representation

This formulation expresses the hypothesis that the relation between LS and syntax is structurally isomorphic, that is, structure is projected from the lexicon. This implies that a VP-node must be present in Hungarian syntax. In the following chapter, I will present empirical evidence for this claim.