Abstract

Recent lexicalist analyses of predicates expressed by syntactically independent elements, raise central questions concerning the domain in which such complex predicates are composed. Should they be composed in phrase structure or within the lexicon? We will argue that a demonstration of syntactic separability for pieces of complex predicates is independent of whether such predicates should be viewed as being composed in the lexicon or phrase structure. We examine Hungarian complex predicates consisting of a syntactically separable preverb, inflectable for person/number features, and a verbal stem. We suggest that the interpretation of the person/number features as oblique pronominals governed by the complex predicate provides an argument for the assumption that predicates expressed by a several syntactically independent elements are better analyzed as associated with lexical representations than as composed in phrase structure.

The Composition of (Dis)Continuous Predicates: Lexical or Syntactic?[1]

Farrell Ackerman and Gert Webelhuth

1. Pieces of predicates with syntactic independence

The past few years have witnessed increased attention within lexicalist frameworks such as Lexical Functional Grammar (LFG) and Head Driven Phrase Structure Grammar (HPSG) to the analysis of predicates expressed by syntactically independent pieces. The recognition of such phenomena, primarily represented in the literature by analytic or periphrastic causatives, raises central questions concerning the domain in which such complex predicates are composed. Given evidence for syntactic separability of the pieces of these predicates, is it compelling to assume that they must be formed in phrasal structure? If so, then the answer to the question posed in the title is that predicate composition is syntactic. Alternatively, are there reasons to argue that such compositions are still lexical, despite the obvious fact of surface independence for the pieces used to expressed these predicates? In related fashion, what would be the basis for a lexicalist perspective on such compositions and in what way would such a lexicalist approach differ from standard lexicalism?

In order to understand the nature of these questions consider the following. It is well-known that Russian contains morphological predicates consisting of a prefix and a verbal stem. These predicates are standardly analyzed as morphophonologically integrated units representing atomic entities with respect to syntactic structure. We will refer to them as synthetic forms of predicates. An example is provided in (1) containing the prefix *ob* 'around': this prefix correlates with an increase in transitivity for the verbal stem yielding the direct object argument 'lake'.

(1) guljajuschie pary **ob**xodjat ozero
strolling pairs around-go-3/pl lake-ACC
`The strolling couples walk around the lake'))

As is to be expected given the morphological status of this wordform, predicates such as these have clear derivatives, both nominal (2) and adjectival (3), related to them:

(2) **ob**xod N 'round' (as in 'make the rounds')
(3) **ob**xodnyj A 'roundabout'

As in Russian, Hungarian has predicates where a preverbal (PV) element modifies certain lexical properties associated with the verbal stem. For example, in (4) we see an instance where the preverb *bele* 'into' correlates with an alteration of both the case

government pattern and the meaning associated with the verbal stem *szol* 'speak, say, talk': whereas *szol* is a one-place predicate, *beleszol* is a two-place predicate which governs the illative case for its oblique complement.

(4) András **bele**szolt a vitá**ba**
András into spoke the dispute-ILL
`András intervened in the dispute'

Once again, as in Russian, the predicate appears to have a morphological status, serving as a base for derivational processes such as nominalization. In the present instance, the verb *beleszol* 'intervene' corresponds to the derived nominals *beleszolás* 'intervention' and *bele nem szolás* `non-intervention'.

These obvious parallelisms between the predicates in Russian and Hungarian clearly suggest a uniform analysis. Such an analysis would appear to be compatible with standard lexical treatments, since, as presented thus far, we seem to be dealing with constructions that not only alter meaning, function assignments and determine case government[2], but are similar from a morphological perspective as well, since the predicates in both languages appear to be complex morphological compositions. Lexicalist accounts are well-suited to address such alterations of functional-semantic information when they are encoded by morphological entities. On the other hand, there is a property characteristic of the Hungarian predicates that distinguishes them from their Russian analogs: in Hungarian the preverb and the verb can function as independent elements in phrase structure. This independence is exemplified in (5) where the presence of the sentential negation element *nem* 'no' immediately to the left of the verbal stem correlates with the postposing of the preverb:

(5) András *nem* szolt **bele** a vitá**ba**
András not spoke into the dispute-ILL)
`András didn't intervene in the dispute'

Formations whose pieces exhibit this sort of syntactic independence are often referred to as phrasal predicates given their analytic or periphrastic expression.

Estonian, like Hungarian, possesses phrasal predicates. In (6) the preverb *ära* 'away' is associated with the predicate *ära ostma* 'corrupt, suborn'. This predicate is based on the simple verb stem *ostma* 'buy, purchase'. The preverb appears discontinuous from the verbal stem at the end of the clause in (6).

(6) mees ostab ta sõbra ära
man buy-3sg his friend-GEN away)
`The man is bribing his friend'

Predicates consisting of a separable preverb and a verbal stem can serve as bases for derivational operations. The following deverbal adjectival and nominal forms related to *ära ostma* 'corrupt, suborn' typify this possibility:

(7)

äraostmatu	A	'incorruptible	äraostmatus	N	'incorruptibility
äraostetav	A	'venal, corrupt'	äraostetavus	N	'venality'

Finally, the phrasal predicates of Hungarian and Estonian resemble in relevant ways one type of German predicate, namely, predicates containing so-called separable particles.[3] An example is provided below containing the predicate *abrufen* 'call up'.

(8) weil wir die Informationen jederzeit *ab rufen* können
because we the information always up call can
`because we can call up the information at any time'

(9) Wir *rufen* die Informationen jederzeit *ab,*
we call the information always up
`We call up the information at any time'

As can be seen, the separable preverb *ab* appears at the end of the finite matrix clause in (9): the verbal stem and preverb are discontinuous in the syntax. As in Hungarian and Estonian, German phrasal predicates may serve as bases for derivational operations. This is exemplified by the possibility for a phrasal predicate to participate in adjective formation with the suffix *-bar* 'able' as in (10):

(10) weil die Informationen jederzeit *ab-ruf-bar* sind,
because the,information always up-call-able are,
`because the information is obtainable at any time'

The predicates in Russian, Hungarian, Estonian, and German all: (i) exhibit lexical effects, i.e., the preverb-V may differ from the verb stem with respect to adicity, semantics, case government, (and grammatical functions) and (ii) exhibit morphological effects, i.e., the preverb and V together constitute a morphological base for derivational and inflectional operations. On the other hand, Hungarian, Estonian, and German differ from Russian in allowing the preverb and verb to exhibit syntactic independence. The existence of phrasal predicates with the profile exhibited by Hungarian, Estonian, and German is widespread cross-linguistically and has elicited the following characterization by Watkins with respect to Indo-European (1964:1037):

> PV V compositions constitute "single semantic words", comparable to simple lexical items; yet they permit *tmesis*, or syntactic separation, suggesting that internal parts are independent syntactic entities.

Phrasal predicates represent an "analytic paradox" with respect to standard assumptions of lexicalism [cf. Nash (1982)]. In particular, their semantic and morphological unithood conflicts with their syntactic separability if the lexicon is interpreted as the source for words employed as syntactic atoms and the syntax as a system for combining and ordering them.

From a cross-linguistic perspective phrasal predicates of the sort illustrated above represent only one type of predicate whose pieces are expressed by syntactically independent elements. For example, there has been an enormous amount of research into causative constructions suggesting that causatives expressed by a single complex wordform, i.e., typified by the Hungarian morphological causative in (11), may exhibit essentially identical semantic effects, grammatical function assignments, case government patterns, etc., as causatives expressed by syntactically separate entities, i.e., typified by the Hungarian periphrastic causative in (12).

(11) a fiú elvonszoltatta Jánost (a hölggyel/a hölgy által)
the boy away-drag-CAUS-3sg/DEF,John-ACC the lady-INSTR/the lady by
`The boy had Janos dragged away (by the lady)'

(12) a fiú, hagyta Jánost elvonszolni (a hölgy által)
the boy let-PAST-3sg/DEF John-ACC away drag (the lady by)
`The boy let Janos be dragged away (by the lady)'

Both (11) and (12) are arguably mono-clausal constructions containing identical *causer* arguments, i.e., `the boy', *patient* arguments, i.e., `John', and optional *causee* arguments, i.e., `the lady'.

Recently there has been a move within lexicalist theories to explain such similarities by positing predicate composition operations which combine certain sorts of information associated with the syntactically separate pieces within phrase structure. This procedure, reminiscent in significant ways of proposals within Government and Binding theory such as Rosen (1989), Baker (1998) among others, is referred to as *predicate composition* by Alsina (1993) and Butt (1995) within the Lexical Functional Grammar framework. On this analysis, the a(-rgument) structures associated with each of the participating predicates combine to create a composite argument structure. This a-structure serves as the basis for assigning grammatical functions to arguments of the complex predicate.

It is important to note that this type of proposal represents a departure from certain long held assumptions concerning the locus for manipulations of lexical semantic information and grammatical function assignment within lexicalist theories. In particular, it

departs from the common assumption that all meaning changing, function changing, valence changing and case-government altering operations are limited to the lexicon.

In section 2 of we examine the nature of the assumptions at issue with respect to such a proposal and will offer an alternative interpretation of lexicalist assumptions. In section 3 we will demonstrate how this alternative conception of lexicalism appears to provide a theoretically satisfying account analysis of Hungarian inflecting preverb and verb combinations, and one which we argue to be preferable to the syntactic composition accounts currently favored among lexicalists.

In general, we will argue that a demonstration of syntactic separability for pieces of complex predicates is independent of whether such predicates should be viewed as being composed in the lexicon or phrase structure. The view of lexicalism defended here will assume, in fact, that predicates expressed by a single syntactic atom as well as predicates expressed by several such atoms are profitably associated with lexical representations. We will forego in the present paper a detailed implementation of these latter assumptions and refer the reader instead to the detailed exposition in Ackerman and Webelhuth (forthcoming).

2. Conceptions of Lexicalism

In our view lexicalism may be regarded as a cluster concept admitting of some gradient among different approaches. In this section we identify three central proto-concepts associated with lexicalism. This will help us to characterize the nature of lexicalism propounded by several different recent approaches depending on which of the principles are recognized in the particular theory. In addition, we can compare the views developed in the present article to these other conceptions of lexicalism. The table in (13) provides an overview of our comparison and the following text explains the meanings of the principles and the values that we have assigned to the cells:

(13) Overview of Lexicalism

Theory	Lexical Adicity	Morphological Integrity	Morphological Expression
Classical LFG and HPSG	yes	yes	Principle
Recent LFG and HPSG	**no**	yes	Principle
This article	yes	yes	**Preference**

We will call the first lexicalist proto-principle *Lexical Adicity* since it relates to the locus for establishing a set of adicity structures for lexical items:

(14) Lexical Adicity

> The adicity of a lexical item is lexically fully determined and cannot be altered by items of the syntactic context in which it appears.

Lexical adicity is intended to cover three different types of information associated with a lexical item: the number and type of its semantic arguments, the number and type of its functional arguments, and the number and grammatical categories of its phrase-structural dependents. For a verb such as the Hungarian predicate *vonszol* `drag', lexical adicity would require that its semantic arguments "dragger" and "dragee", its functional arguments "subject" and "object", and its categorial arguments "NP[nom]" and NP[acc] already be specified in its lexical entry. The causative lexical entry (or predicate) in (11) based on *vonszol* "hit", specifically *elvonszoltat* `make drag', likewise would be lexically completely specified for semantic, functional, and categorial selection, because (14) reserves the power of specifying these selectional properties for the lexicon and expressly withholds this privilege from the mechanisms applying in the syntactic component.

As the table indicates, classical LFG and HPSG both incorporated lexical adicity. In the context of the theories presented in Bresnan (1982) or Pollard and Sag (1987) the selectional properties of lexical items were completely determined in the lexicon and all changes in the meaning of a predicate or its selectional properties were achieved in the lexicon (via lexical rules) and were independent of the syntactic context into which the lexical entry was inserted.

Recent LFG and HPSG analyses of complex predicate phenomena, however, extend the privilege of creating new argument structures from the lexicon to the syntax, in direct violation of *Lexical Adicity.* In the case of LFG, Alsina (1993, iv, v, 280) admits "partially specified predicates" whose adicity is only fixed in the syntactic component, as can be inferred from the two quotes below:[4]

> The operations that affect the way that arguments are overtly expressed are assumed to be operations on the argument structure of a predicate and are treated as partially specified predicates that must compose with other predicates to yield fully specified predicates. Thus, predicate composition is responsible for operations such as passivization, causativization, applicativization, etc.

> Most work within LFG, and other lexicalist theories, has assumed that predicate composition, or the equivalent notion in each particular theory, can only take place in the lexicon. However, the evidence indicates that causative (and other) complex predicates in Romance are not derived in the lexicon because the two verbs that compose the complex predicates do not constitute a word. If the lexicon is the word formation module of the grammar and words are the terminal nodes of the c-structure, we have to conclude that causative constructions in Romance contain two words that jointly determine the predicate of the clause. This forces us to design a theory that allows predicate composition to result not only from combining morphemes in the lexicon, but also from combining words and phrases in the syntax. In what follows, I will first present evidence that the causative complex predicate in Romance does not correspond to one word (a morphological unit) or even one single X^0 or terminal node in the syntax, and that it is, therefore not formed in the lexicon; and I will then indicate the necessary assumptions for an LFG theory to allow predicate composition in the syntax.

Within HPSG, the highly influential proposal of Hinrichs and Nakazawa (1989) allows lexical entries to subcategorize for another lexical entry as a complement. As a consequence, the selecting lexical entry may inherit some or all of the selectional properties of that complement. This yields a configuration where a selector with an initially underspecified argument structure comes to have a fully specified argument structure. Thus, an auxiliary that selects for a main verb complement and inherits all of that complement's arguments will have a different number of arguments depending on whether the embedded complement has zero, one, two, or three arguments. Since the identity of the verb that serves as the complement to the auxiliary will only be known once the two verbs appear together in phrase structure, the argument structure of the auxiliary will be finally specified only in the syntactic component as a function of the syntactic context in which the auxiliary appears. This is in clear violation of the principle of *Lexical Adicity.*

LFG and HPSG thus have undergone a conceptual transformation in their recent history in that both theories have reset the boundaries between the applicability of lexical and syntactic mechanisms in favor of the syntax: whereas previous versions of both approaches gave certain analytical privileges to the lexicon and withheld them from the syntax, the recent versions of these theories allow the syntax to move further into the territory once held exclusively by the lexicon.

In this connection it is important to appreciate that the empirical motivation for this relative loss of distinction on the part of the lexicon is precisely the set of phenomena dealing with analytically expressed clausal heads (i.e. predicates). Alsina (1993), Butt (1995), and Hinrichs & Nakazawa (1989) all motivate the need for the creation of new argument structures in the syntax on the basis of constructions involving a combination of two verbs which jointly define the semantic, functional, and categorial properties of a clause, e.g. a combination of a causative verb and a main verb or a combination of an auxiliary and a main verb.

As the entry in the final row of table (13) indicates, the theory of predicates advanced in this article retains the strongly lexicalist position of classical LFG and HPSG: the lexicon and *only the lexicon* has the privilege of specifying the properties that make up the adicity of a phrase-projecting head. We believe that it is the wrong theoretical choice to weaken the influence of the lexicon relative to the syntax in the face of analytically expressed predicates and - as will be stated shortly - instead take the position that this problem is most effectively solved by realigning the relative influences of the lexicon and the syntax in the other direction. In other words, the type of theory assumed in the present article will force the syntax to cede some further analytical ground to the lexicon and hence in this respect is an even more strongly lexicalist theory than classical LFG and HPSG.

Our second proto-principle of lexicalism deals with the relationship between the lexical component and morphology:

(15) Morphological Integrity

> Syntactic mechanisms neither make reference to the daughters of morphological words nor can they create new morphological words in constituent structure.

In the words of Di Sciullo and Williams (1987), *Morphological Integrity* creates a "bottle neck" represented by morphological words: the sole morphological information that syntax can access is the morphology of the root node of a morphological constituent structure tree. Syntax cannot "look" lower in the tree at the root's daughter constituents. Bresnan and Mchombo (1995) present this point as follows (note that these authors prefer the term *Lexical Integrity* to the somewhat more specific *Morphological Integrity):* [5]

> A fundamental generalization that morphologists have traditionally maintained is the *lexical integrity principle,* that words are built out of different structural elements and by different principles of composition than syntactic phrases. Specifically, the morphological constituents of words are lexical and sublexical categories - stems and affixes - while the syntactic constituents of phrases have words as the minimal, unanalyzable units; and syntactic ordering principles do not apply to morphemic structures ... it has been hypothesized that *the lexical integrity principle holds of the morphemic structure of words, independently of their prosodic or functional structure.*

We take *Morphological Integrity* to mean that syntax and morphology are separate but interacting domains of grammar. Syntax, interpreted as phrasal structure, can neither "look into" morphological words to see internal structure nor can it create new morphological words. The lexicon is not subject to either of these two constraints and hence has a more privileged relation to morphology than the syntax.

Each of the theories compared in our overview table (13) claims this morphological privilege of the lexicon over the syntax and in so doing they all differ from other theories

that do allow morphological and syntactic operations to be intermixed, e.g. many versions of Government and Binding Theory and classical Montague Grammar.

The third and final diagnostic entering into an explication of lexicalism will be referred to as *Morphological Expression:*

(16) Morphological Expression

> Lexical expressions are uniformly expressed as single synthetic (syntactically atomic) word forms.

The concept of morphological expression, we believe, has been mistakenly conflated with morphological integrity as characterized above. Specifically, whereas morphological integrity constrains syntactic operations from creating morphological wordforms, morphological expression concerns assumptions about the surface means by which lexical representations are expressed. LFG and HPSG have traditionally held the lexicon to the strict requirement that each lexical representation be expressed by at most one single morphophonologically integrated word form[6]. This requirement privileges the syntax to create all collocations that consist of more than one morphological piece, even if the ensemble of morphological pieces behaves as one functional-semantic unit with one argument structure, e.g. the analytical causatives discussed in Alsina (1993) and the auxiliary-verb combinations discussed in Hinrichs and Nakazawa (1989). It is precisely this required connection between clausal heads inserted from the lexicon and single morphological surface forms that leads these authors to abandon the restriction against the formation of new argument structures in the syntax as was discussed in connection with the principle of *Lexical Adicity.*

There is thus conceptual tension between *Lexical Adicity* and *Morphological Expression,* and this tension becomes most obvious in the treatment of analytically expressed clausal heads. Classical LFG and HPSG maintained both principles but were unable to provide optimal analyses of these types of heads. Two types of responses are possible to this state of affairs and both involve a realignment of the relative privileges of the lexicon and the syntactic component, albeit in opposite directions. Either one considers it of paramount importance to retain the morphological restrictions of the lexicon vis-a-vis the syntax: then one is led to create analytically expressed clausal heads in the syntax by allowing phrase-structural operations to invade into the previously exclusively lexical domain of the formation of new argument structures. This leads to the departure from classical lexicalism that is represented by works such as Alsina (1993) and Hinrichs and Nakazawa (1989). Accordingly, lexicalism is in a weaker position relative to the syntax in

recent LFG and HPSG compared to the classical versions of these theories (see the first and second rows in (14)).

Alternatively, if one considers *Lexical Acidity,* i.e. the exclusive privilege of the lexicon to create and manipulate the functional-semantic information associated with clausal heads, to be the conceptual core of lexicalism, one can still maintain a principled role for *Morphological Expression* : interpreting it as a markedness preference for the encoding of lexical representations strengthens the relative analytical role of the lexicon vis-a-vis the syntax. Whereas classical lexicalism allowed the syntax to deal with collocations without joint morphological status and withheld this option from the lexicon, *Morphological Expression* as a markedness principle makes the syntax only the preferred locus of composition for analytically expressed elements but extends this option to the lexicon as a marked choice. The unmarked choice of expression for a lexical item is, of course, the sort of integrated morphophonological entity which motivates *Lexical Integrity.*

To sum up our discussion of lexicalism as a cluster concept: this article takes the view that the data from predicates expressed by syntactically independent elements do not warrant abandoning what we take to be foundational principles of lexicalism, in particular the principle we called *Lexical Adicity* which prevents the syntactic component from creating new argument structures. The argument developed in this article is guided by the conviction that this functional-semantic component of lexicalism should only be abandoned if the puzzles created by (complex) predicates prove to be thoroughly incommensurable with all defensible implementations of this view. From a more positive perspective, we will demonstrate that adherence to these functional-semantic principles raises important questions and yields important results. Accordingly, our overall view can perhaps best be characterized as follows:

(17) The Primacy of Function over Form

> Lexicalism is first and foremost a hypothesis about functional-semantic information and secondarily a hypothesis about form.

Given this general perspective on lexicalism, we are led to postulate the profile of principles in the last line of the overview table of lexicalism. This proposal can be summed up for easy reference as follows:

(18) Assumptions of the Present Article

- Only lexical but not syntactic rules can create new argument structures *(Lexical Adicity).*
- Only lexical but not syntactic rules can create or analyze morphological words *(Morphological Integrity).*
- Lexical representations are preferably expressed by single synthetic word forms but can also be expressed by combinations of words without joint morphological status *(Morphological Expression).* [7]

In effect these assumptions lead to two contending interpretations of predicates whose pieces exhibit syntactic independence. On the prevailing account, the pieces only interact with one another when they co-occur in phrase structure: information associated with the pieces can only be composed when these pieces co-occur. The alternative argued for below is to jointly associate the independent pieces with a single lexical representation. In section 3 we will explore a domain in which it appears useful to be able to appeal to lexical representations associated with syntactically independent elements.

3. Morphological Blocking vs. Lexical Blocking

As observed in Poser (1992), a phenomenon called *blocking* has been a traditional source of explanation within morphology. He characterizes this phenomenon in the following way: "the existence of one form prevents the creation and use of another form that would otherwise be expected to occur" [Poser (1992:11)]. For example, Aronoff (1976) appeals to blocking in order to account for certain restrictions on nominalization in English: whereas adjectives bearing the suffix *-ous* typically have nominal counterparts with the suffix *-ity* e.g. *curious* and *curiosity*, adjectives without such nominal counterparts appear related to nominals that are listed and thereby 'block' the formation of the derived and expected form, e.g. *glorious* and *glory* , but not **gloriosity.*

Similarly, analyses formulated within lexical phonology/morphology as developed in Kiparsky (1982) and K.P. Mohanan (1986) have employed blocking to account for numerous phenomena. For example, Gordan (1985) has appealed to level ordering within lexical morphology in order to explain the apparent predisposition of children acquiring English to permit irregular plural forms as left members of synthetic compounds. In particular, he observed that children are prone to produce (in experimental settings) compounds such as 'teeth-eater' containing the irregular plural 'teeth'. In contrast they produce such forms as 'bead-eater' rather than *'beads-eater': these compounds contain left members which have regular plural forms. Given the level ordered assumptions that irregulars are present at a level prior to compounding and regular forms are present at a

level after compounding, the observed behaviors are argued to follow: the irregular forms can feed compounding and therefore are possible as left members of compounds, while regular forms cannot feed compounding since they follow this operation. In sum, the organization of the morphological component into levels where irregular forms serve to block regular ones can be used to explain certain subtle constraints on compounding.

Noting that 'blocking' has been restricted to relations between 'lexical items' and interpreting 'lexical items' in conformity with Morphological Integrity and Morphological Expression as previously discussed, Poser (1992) presents several phenomena where 'lexical items' appear to block certain phrasal constructs. We will discuss one of the phenomena mentioned by Poser since it is particularly relevant to the central issue of this paper: periphrastic verbs in Japanese.

We will see that if lexical representations of predicates are separated from their morphological expressions, as argued for in the present work, then we are confronted by what can be referred to as 'lexical blocking'. That is, one lexical representation blocks the appearance of another lexical representation and does so in accordance with the markedness principle of expression for lexical representations previously discussed. In other words, the approach to predicates offered in the present work extends to account in a straightforward way for instances where single morphological objects can block ensembles of morphological objects, because both sorts of entities are hypothesized to be expressions of lexical representations.

We begin by discussing Poser's example and then turn to two other similar phenomena, specifically, verbal inflection in Irish [following Andrews (1990)] and oblique inflection on preverbs in Hungarian [following Ackerman (1987)].

3.1 "Phrasal" Blocking

Poser (1992) observes that Japanese contains minimal pairs of periphrastic verb forms which are commonly thought of as unincorporated predicates as in (19a) vs. incorporated predicates as in (19b):

(19a) Eigo-no benkyoo-o site-iru
English-gen study-acc doing-be)
`(He) is studying English'

(19b) Eigo-o benkyoo site-iru
English-acc study doing-be)
`(He) is studying English'

Both (19a) and (19b) contain the verbal element *suru* 'do.' On the other hand, this verb is preceded by an accusatively marked nominal object in the unincorporated variant, but in the

incorporated variant it is preceded by an unmarked nominal. Poser argues that despite the fact that the incorporated variant is frequently treated as a single wordform, there is evidence to suggest that it is actually phrasal, like the unincorporated type. He provides the following evidence in support of this claim: (1992:112)

(a) Periphrastic verbs are accented like phrases rather than like any other sort of verb,
(b) Reduplication affects only the *suru* component of the periphrastic,
(c) Periphrastic verbs do not undergo even highly productive lexical nominalizations,
(d) Sentence-internally periphrastics are analyzable into the nominal verbal portions, in that the nominal may be omitted in whether constructions, which require repetition of the verb,
(e) It is possible to Right Node Raise the *suru* portion alone,
(f) It is possible to delete the verbal noun in the second conjunct of a pair of conjoined sentences,
(g) Periphrastics are analyzable across sentence-boundary in that the nominal part may be omitted in *too*-clauses, in which the verb of the first sentence is repeated in the second sentence,
(h) Periphrastic verbs are analyzable at the discourse level across speakers into the nominal and *suru*, in that the nominal part may be omitted in responses to yes-no questions.

Having demonstrated that the incorporated type of periphrastic verb is phrasal, Poser then goes on to show that these forms are blocked by simple verb forms. This argument is based on the interaction of deverbal noun formation and periphrastic predicate formation. In particular, he notes that Japanese possesses a productive deverbal noun formation process yielding nominals such as *iri* 'parching' and *mamori* 'protecting' from *ir* 'parch' and *mamor* 'protect'. However, despite the fact that it is possible to create these deverbal nouns, the language does not permit them to be used in periphrastic predicate formation, i.e. **iri suru* 'parch' or **mamori suru* 'protect'. Following Kageyama (1982), he suggests that the prohibition against such periphrastic forms is directly related to the existence of simple verbs which block them, i.e. *ir* 'parch' and *mamor* 'protect.' He concludes that "... this seems to be a case in which lexical items block a phrasal construction" (1992:119).

As suggested previously, a lexical representation of predicates which allows for several sorts of expression types renders possible precisely the sort of analysis which makes sense of such effects. On the account developed here we are confronted with "lexical" blocking, where the blocking of synthetic morphological objects by synthetic morphological objects is simply one common type of blocking. Since, by hypothesis, there are other expression types for lexical representations, specifically phrasal expressions, it is to be expected that certain synthetic expressions might block certain phrasal expressions as long as each is associated with a lexical representation. That is, a lexical representation with a certain information content and synthetic expression tends to block the use of an analytic expression of a lexical representation which could convey the same information. Andrews

(1990) represents a particularly perspicuous formulation of this effect which he calls The Morphological Blocking Principle.

3.2 The Morphological Blocking Principle

Andrews (1990) proposes a morphological blocking principle (MBP) in order to explain, among other phenomena, the following sort of distributions in Irish verbal inflection:

(20a)	Chuirfinn		isteach,	ar	an	phos	sin
	put-CONDIT-S1		in	on	the	job	that
	`I would apply for that job'						
(20b)	*Chuirfeadh	mé	isteach	ar	an	phost	sin
	put-CONDIT	I	in	on	the	job	that
	`I would apply for that job'						

The verb in (20a) is an inflected form which specifies features for the subject argument of the verb meaning 'apply:' in particular, it bears first person singular features. In contrast, the verb in (20b) is uninflected and therefore does not morphologically express any features of the subject: the subject features of the predicate meaning 'apply' are supplied by the independent first person subject pronoun appearing in constituent structure. The ungrammaticality of (20b) suggests that though in principle one might think it possible for an independent pronoun to satisfy the subject complement requirement of the predicate, it turns out it cannot.

Andrews, following an observation by McCloskey and Hale (1984), develops the hypothesis that, for purposes of satisfying the functional requirements of predicators, uninflected verbal forms [i.e. the verbal form in (20b)] are dispreferred when inflected forms are available. He formalizes this hypothesis as the Morphological Blocking Principle:

> Suppose that the structure S has a preterminal node P occupied by a lexical item l_1, and there is another lexical item l_2 such that the f-structure determined by the lexical entry of l_1 properly subsumes that determined by the lexical entry of l_2, and that of l_2 subsumes the f-structure associated with P in S (the complete structure, after all unifications have been carried out.) Then S is blocked.

The effect of this principle is as follows: when there are alternative ways of satisfying some requirement of a lexical item, e.g. finding a value for a grammatical function such as SUBJ, then satisfaction via an inflected form specifying the relevant value is to be preferred over a form which does not have such a specification and which would, consequently, require satisfaction of the requirement in constituent structure. As applied to the examples in (20), this means that the verb 'apply' requires a value for its subject argument: since there is an inflected form which can determine a value for this function, it is to be preferred over the

uninflected form in (20b) which could only determine a value for this function by an independent pronoun in constituent structure. Note that MBP is formulated in terms of the relative content of related lexical items, the basic idea being that an item specified for some property blocks the use of a related item not specified for that property.

On the assumption that the MBP is a principled way to distinguish between alternative expressions of related lexical items, it yields empirical predictions that can help to distinguish between two views of predicate composition: specifically, it can help to distinguish between the view that predicates consisting of demonstrably independent syntactic elements are composed in the syntax - the view currently adopted in one form or another among practitioners of LFG and HPSG, as outlined in section 2 - in contrast to the view developed here, namely, that predicate composition is an operation among lexical representations associable with different expression types.

On a syntactic composition account, if a language contains predicates with demonstrably separate syntactic pieces, then the pieces should not jointly exhibit a lexical blocking effect: this prediction follows since on the syntactic account predicate composition is between two independent elements in constituent structure. From the perspective of the Irish data presently previously, there would be, in effect, two syntactic ways of satisfying the requirements of the predicate: MBP, however, does not account for preferred options among two alternative syntactic ways of satisfying predicate requirements. In contrast, if surface expression is separated from the lexical status of predicates, as on the present account, then such data is precisely what is to be expected: syntactically independent pieces of a predicate could exhibit blocking effects, since they are simply a particular expression type of lexical representations and illustrate the operations of MBP interpreted as a Lexical Blocking Principle (LBP). Accordingly, a more specified lexical representation blocks the use of a less specified related lexical representation. In the next section we examine one type of phenomenon from Hungarian which bears on this prediction.[8]

3.3 Against Predicate Composition in Syntax

Our interpretation of Andrew's MPB as a lexical blocking principle as presented above restricts a less highly specified lexical entity from satisfying the requirements of a predicate if there is a more highly specified synthetically expressed lexical entity which expresses the requisite information. In Irish we saw that both the inflected (i.e. more highly specified) form and the uninflected (i.e. less specified) form were synthetic morphological objects: that is, both of the relevant forms were syntactic atoms. They

differed with respect to whether an independent element in constituent structure could satisfy the requirements of the predicate: the issue was not whether the predicate itself is expressed synthetically or analytically/periphrastically, but rather whether its requirements are satisfiable internal to the synthetic form or external to it. In contrast to the Irish examples, the Japanese data exemplify instances where a synthetic verbal form blocks the existence (or use) of an analytic verbal form: here blocking concerns the preference for synthetic verbal expressions over analytic ones, but argument satisfaction is not at issue.

In the present section we examine an instance where the predicate can be expressed by syntactically independent elements, thus analogous to the sorts of periphrastic predicates in Japanese, while also showing the sorts of argument satisfaction properties reminiscent of Irish. We argue that the phenomenon of oblique argument incorporation in Hungarian is one instance of empirical data which serves to distinguish lexical vs. syntactic composition of predicates, thereby bearing on the general prediction described above.

Consider the following data from Hungarian containing the transitive predicate *szeret* 'love' in (21) and the related two place non-transitive predicate *'bele szeret*: ' **'**fall in love**'** composed of the verbal stem *szeret* 'love' and the preverb *bele* 'into' in (22).

(21) a fiú szerette a lányt/*lányba
the boy loved the girl-ACC/girl-ILL
`The boy loved the girl'

(22) a fiú bele szeretett a *lányt/lányba
the boy PV loved the girl-ACC/girl-ILL
`The boy fell in love with the girl'

The predicate containing the preverb *bele* and a verbal stem in (21) is representative of a class of predicates which differ from the simple verbal stem with respect to meaning, the grammatical function status of arguments and case government. In addition, the preverb and the verbal stem are separable under certain syntactic conditions. For example, preverbs are postposed in contexts of sentential negation as illustrated in (22) where the negative element *nem* immediately precedes the verbal stem:

(23) a fiú **nem** szeretett **bele** a *lányt/lányba
the boy not loved PV the girl-ACC/girl-ILL
`The boy didn't fall in love with the girl'

The syntactic independence of the verbal stem and the preverbal element in conjunction with the clear differences concerning function assignment etc. viz. the simple predicate makes such constructions natural candidates for an analysis in terms of syntactic composition. In this connection it is important to consider the manner in which the oblique argument requirement of the complex predicate can and cannot be satisfied for pronominals

(for detailed discussion see Ackerman (1987, 1990) as well as an alternative account in E. Kiss (forthcoming)). This contrast is illustrated in (24).

(24a) a fiú belé**(je)** szeretett
the boy PV-3sg loved
`The boy fell in love with him/her'

(24b) *a fiú bele szeretett beléje
the boy PV loved 3-sg-ILL
`The boy fell in love with him/her'))

(24c) *a fiú belé**je** szeretett a lányba,
the boy PV-3sg ,loved girl-ILL
`The boy fell in love with the girl'

In (24a) the preverb bears a marker from the possessive paradigm which functions as an oblique pronominal, while the uninflected preverb in (24b) cannot co-occur with an independent pronominal in the appropriately case governed form. Finally, the inflected preverb in (24c) cannot co-occur with the lexical NP. In general this is the pattern for pronominal satisfaction for this subtype of predicate: the preverb bears a marker from the possessive paradigm indicating the person/number features of the pronominal argument, while an uninflected preverb co-occurs with an independent lexical NP in order to satisfy the oblique requirement non-pronominally (see example (22)).[9] This distribution, accordingly, recalls the Irish facts accounted for by the MBP. However, there is a salient difference between Hungarian oblique incorporated pronouns and their Irish subject function analogs: whereas the Irish data involved inflected vs. uninflected forms of synthetic predicates, the Hungarian data involve inflected vs. uninflected syntactically independent pieces of a predicate.

The obvious question is this: if the Hungarian predicate is composed in the syntax, how could it exhibit the observed blocking effects? On the assumption that predicate composition is syntactic and that the satisfaction of argument requirements by an independent pronoun is syntactic, it would seem that a syntactic operation, i.e. predicate composition, blocks the syntactic satisfaction of an argument requirement. In what sense could the observed blocking facts be assimilated to the class of effects which the MBP (in our view, better construed as LBP) was designed to explain?

One way of answering this question goes as follows. In lexicalist frameworks where syntactic composition is advanced the lexicon contains fully inflected forms. Given this, the inflected form could be taken directly from the lexicon [see, for example, Niño (1995), Nordlinger (1995)]. Obviously, the inflected form of the preverb contains more information than the uninflected form of the same preverb. In this sense there is potential

competition concerning argument satisfaction between a morphologically more marked form and a less marked form. If the inflected form, i.e. the more highly specified form, is preferred over the uninflected form, i.e. the less specified form, in order to satisfy argument requirements of the predicate, then this would conform to the expectations of the MBP. In other words, a syntactic composition account might achieve the MPB effects by positing a lexically provided contrast between inflected vs. uninflected forms. In sum, one might argue that the inflected preverb itself is more highly specified than its uninflected form and that, given a pronominal interpretation of the inflected form, the MBP applies to the paradigmatic contrast represented by these lexical forms.

An analysis of the preceding type is based on the assumption that the inflected preverb satisfies the argument requirements of a co-occuring predicates. Now the question arises, of course, as to what predicate is being satisfied such that the inflected preverb can serve to satsify it. Recall that in example (21) the verb *szeret* means 'love' and requires a subject and an object argument. The inflected preverb cannot be satisfying the requirement of this verb, since this verb doesn't permit, let alone require, an oblique argument: this is evident by the fact that the variant of (21) with an ILLATIVE complement is ungrammatical. Neither a lexical NP in the ILLATIVE case nor an inflected preverb can co-occur with the simple predicate. Rather the inflected form is satisfying the argument requirements of the predicate *bele szeret* 'fall in love' consisting of a separable preverb and a verbal stem. An inflected form can only be taken from the lexicon if it is assumed that the relevant predicate is e.g. *bele szeret*.. This is evident from the distributions in (22) and (24).

In other words, a solution that posits that an oblique co-occurs with the simple verbal stem is poorly motivated, though of course it is possible to stipulate that when an inflected preverb co-occurs with certain verbal stems the two should be interpreted as a complex predicate, i.e. that syntactic composition is obligatory. Given such a stipulation, then, it might be argued that a morphologically more specified form blocks the syntactic satisfaction of the argument requirements of the complex predicate for pronominal satisfaction, since the preverb is presumably inflected in the lexicon. This gives the appearance that syntactic composition is adequate to achieve the desired patterns.

On the other hand, such a solution is somewhat paradoxical: the interpretation of the inflected preverb as forming a complex predicate with the verb is only done when one knows antecedently that an uninflected preverb and the verbal stem form a complex predicate, exhibiting properties other than those shown by the simple verbal stem. In other words, one only composes the relevant elements in syntax because it is known that they form a functional-semantic unit: syntactic composition thus appears to represent an operation that achieves correct results but applies precisely when the syntactically

independent elements are known to possess a lexical representation, as suggested on our analysis.

The problem of when syntactic composition must apply is further exacerbated by the fact that Hungarian has dozens of different types of preverb and verb compositions whose meanings, argument inventories, function assignments and case government requirements span the scale from purely predictable to idiosyncratic: some properties are sometimes predictable, while others may not be. In the type of composition examined above, for example, one must know that the combination of *bele* 'into' and *szeret* ' love' means 'fall in love' : that it requires an oblique ILLATIVE complement, however, follows from the presence of the preverb *bele*. Similarly, one must know the meaning of the behaviorally identical formation *bele köt* 'quarrel', containing the transitive verbal stem *köt* 'bind: this predicate possesses an idiosyncratic meaning, but its case marking and oblique argument requirement are fully predictable, given the presence of the preverb *bele.* The class of cases of which *beleszeret* is a member can be characterized as *causal predicates.* This class exhibits roughly the following properties: (i) the verbal base denotes a psychological or physical state, (ii) the verbal stem co-occurs with the preverb *bele,* and (iii) the predicate governs the ILLATIVE case for its oblique argument. In addition, there are morphological restrictions on the form of the verbal base: it cannot contain the transitivizing suffix *-ít* or the causative suffix *-tat*.. For example, whereas the causal predicates *belevakul* 'become blind from' and *beleun* 'become bored from' exist, the corresponding forms **belevakít* 'make blind from' and **beleuntat* 'make bored from' do not.

In summary, a syntactic composition account would have to attribute to syntactic composition operations the possibility of yielding all types of predicates irrespective of their degree of regularity: such operations could create forms which exhibit various types of irregularity (from semantic idiosyncrasy to odd case government patterns). If it is argued that irregular and partly regular forms should be listed, then given the behavioral identity between the regular, partly regular, and irregular forms, it would be unclear why the representational apparatus needed for the irregular and partly irregular forms could not be employed for the fully regular ones. In other words, there would appear to be no need for syntactic composition at all, since it replicates what must already be done lexically. This, of course, would follow from usual considerations of parsimony.

The types of lexical representations compatible with the present account are designed to capture the notion of hierarchical relatedness assumed in Andrew's statement of the MPB. They are therefore appropriate objects for the operation of the MBP. In particular, lexical adacity as stated previously, requires the information associated with skeletal clause nuclei (in particular, all the grammatical function requirements of a clause) as

well as inflectional information (following the Strong Lexicalist Hypothesis) to be encoded in lexical representations, independent of whether a predicate is expressed by one or more syntactic atoms. Given this assumption, the force of the MBP can be maintained: that is, it is a principle that provides a motivation for choosing between alternative expressions of related lexical representations. On the present analysis an inflected form is licensed to occur quite simply as a function of the existence of a lexical representation for the relevant item which has as one of its exponents the uninflected form of the preverb in composition with the verbal stem. In other words, the inflected form is interpretable as a more highly specified form of the uninflected form which is used to express the functional-semantic unit encoded in a lexical representation: it is the grammatical function requirements associated with this lexical representation that is relevant for calculating blocking effects. Since the contrast is between different surface realizations of a single lexical item, this distribution comports with expectations based on the LBP.

In conclusion, it appears that Hungarian may provide us with empirical evidence bearing on the desirability of a lexical vs. syntactic account of predicate formation. We have argued that a conception of lexicalism in which functional-semantic considerations are pre-eminent entails that certain lexical representations will of necessity be expressed by multiple morphological elements in syntax. This is interpretable as marked expression for lexical representations. If such marked expression types are associated with lexical representations, then we can explain why these predicates exhibit the sorts of blocking effects previously attributed to morphological blocking. On the present account, these effects are better interpreted as lexical blocking effects, where predicate associated with lexical representations can participate in blocking irrespective of their surface syntactic encoding.

References

Ackerman, Farrell 1990. The Morphological Blocking Principle and Oblique Pronominal Incorporation in Hungarian. In K. Dziwirek et. al. eds. *Grammatical Relations: A Cross-Theoretical Perspective.* Stanford: CSLI Publications.

Ackerman, Farrell. 1987. Miscreant Morphemes: Phrasal Predicates in Ugric. Doctoral dissertation, University of California, Berkeley.

Ackerman, Farrell and LeSourd, Phil. 1993/1997. Toward a Theory of Phasal Predicates. In P. Sells et. al. eds. *Complex Predicates.* Stanford: CSLI Publishers.

Ackerman, Farrell and Webelhuth, Gert. Forthcoming. *A Theory of Phasal Predicates.* Stanford: CSLI Publishers.

Alsina, Alex. 1993. Predicate Composition: A Theory of Syntactic Function Alternations. Doctoral dissertation, Stanford University.

Andrews, Avery. 1990. Unification and Morphological Blocking. *Natural Language and Linguistic Theory* 8:507-557.

Baker, Marker. 1988. *Incorporation: A Theory of Grammatical Function Changing.* Chicago, Ill.: University of Chicago Press.

Bresnan, Joan. 1982. The passive in lexical theory. In J. Bresnan ed., *The Mental Representation of Grammatical Relations.* Cambridge, Mass.: MIT Press.

Bresnan, Joan, and Sam A Mchombo. 1995. The Lexical Integrity Principle: Evidence from Bantu. *Natural Language and Linguistic Theory* 13.2: 181-254.

Butt, Miriam. 1995. *The Sructure of Complex Predicates in Urdu.* Stanford, Ca.: CSLI Publications.

Di Sciullo, Anna-Maria, and Edwin Williams. 1987. *On the Definition of Word*. Cambridge, Mass.: MIT Press.

Gordan, Peter. 1993. Level-ordering in lexical development. In P. Bloom ed., *Language Acquisition : Core Readings.* Cambridge, Mass.: MIT Press. 485-503.

Hinrichs, Erhard and Tsuneko Nakazawa**.** 1990. Subcategorization and VP structure. In S. Hughes ed., *Proceedings of the Third Symposium on Germanic Linguistics.* Amsterdam: J. Benjamins.

Holloway-King, Tracy. 1995. *Configuring Topic and Focus in Russian.* CSLI Publications.

Kiparsky, Paul. 1982. Explanation in Phonology. Dordrecht: Foris.

E. Kiss, Katalin. Forthcoming. Postposition incorporation in Hungarian. In Caspar de Groot and István Kenesei eds. *Approaches to Hungarian* Vol. 6.

Jackendoff, Ray. 1995. Lexical insertion in a post-minimalist theory of grammar. MS. Brandeis University.

Kageyama, Taroo. 1982. Word formation in Japanese. *Lingua* 57:215-258.

McCloskey and Hale. 1984. On the syntax of person-number marking in Irish. *Natural Language and Linguistic Theory* 4:245-281.

Mohanan. K. P. 1986. *The Theory of Lexical Phonology.* Dordrecht: Kluwer Academic Publishers.

Nash, David. 1982. Verb Roots and Preverbs. *Work Papers of SIL-AAB.* Series A, 6.

Nordlinger, Rachel. 1995. Split tense and imperative mood inflection in Wambaya. MS. Stanford University.

Niño, Maria-Eugenia. 1995. Split inflection and the architecture of universal grammar. MS. Stanford University.

Pollard, Carl and Ivan Sag. 1987. *Information-based Syntax and Semantics.* Stanford Ca.: CSLI Publishers.

Poser, William. 1992. Blocking of phrasal constructions by lexical items. In A. Szabolcsi and I. Sag eds., *Lexical Matters.* Stanford, Ca.: CSLI Publications. 111-130.

Rosen, Sarah. 1990. *Argument Structure and Complex Predicates*. New York: Garland Publications.

Watkins, Calvert. 1964. Preliminaries to the Reconstruction of Indo-European Sentence Structure. In Horace G. Lunt, ed., *Proceedings of the Ninth International Conference of Linguists.* The Hague: Mouton.

---

1 The authors have benefited greatly from collaboration with Phil LeSourd.

2 These properties will be collectively referred to as the `lexical adacity' of predicates in section 2 of this paper.

3 Of course, English particle verb constructions also exemplify this problem. For example, whereas it is possible to say ‘the teacher dressed the boy down' the related nominal is is preferrably ‘the teacher's dressing down of the boy', rather than ‘the teacher's dressing of the boy down.'

4 For a similar view, see Butt (1995, chapter 5 and elsewhere in her book).

5 The view of lexical integrity proposed in Bresnan and Mchombo 1995, i.e., that the leaves of syntactic trees contain fully inflected and derived wordforms and that morphological operations are prohibited from occurring in syntax, is also proposed in Ackerman and LeSourd 1993/1997 with respect to Hungarian.

6 We are, of course, ignoring multi-word idioms in this discussion, since they are generally acknowledged to be listed and do not fall within the purview of this assumption.

7 Familiar accounts of “lexical insertion” deal only with synthetically expressed predicates. On our alternative view the question arises how the parts of an analytic predicate are associated with positions in syntactic structure. [Cf. Jackendoff (1995) for similar considerations concerning lexical entities and lexical insertion]. The details of our proposal are presented in Ackerman and Webelhuth (forthcoming). It provides lexical representations for several sorts of predicates in a unification-based type-theoretic formalism which also addresses the syntactic realization of potentially discontinuous pieces of predicates.

8 See Holloway-King 1995 for an examination of future marking in Russian where it is argued that a lexical entity is argued to block a syntactic phrase.

9 It is worth noting that the base form of inflectable preverbs, as in (22), corresponds to one of the variants of the inflected form for 3sg: this can be seen in the optionality of the 3sg marker indicated in

(24a). There are at a few reasons why it is probably preferable to assume that the prefix *bele* when governing a lexical NP is unmarked, rather than construed as paradigmatic gap for 3sg. First, the absence of an oblique governed argument of predicates containing uninflectable preverbs also is interpretable as conveying a missing 3rd person referent which unspecified for number. For example, the predicate *be ment* `into go' containing the uninflectable preverb *be* can function as a clause meaning `S/he went in (to some known place(s)).' In other words, the interpretation of a missing 3sg argument in (24a) is not dependent on the presence of an unexpressed 3sg possessive marker, but is rather a typical instance of a null complement established by a governing (complex) predicate. Second, and conversely, if the preverb *bele* were construed as containing a 3sg possessive marker, we would have no explanation for the acceptability of (22) and the unacceptability of (24c), since an expressed pronominal element cannot co-occur with a lexical NP in the latter instance. Finally, if the base form were construed as an optionally expressed 3sg element, it is not clear how on would explain in a non-circular manner how the uninflected form could co-occur with plural lexical NPs as in (i):

(i)	*a	fiú	belé(***je)**	szeretett	a lányokba
	the	boy	PV-3sg	loved	girl-PL-ILL
	`The boy fell in love with the girls'				

In sum, it is reasonable to assume that the uninflected base form in (22) is interpretable as not representing a paradigmatic gap.

In connection with the distributions associated with (24), a reviewer observes that the quantifier *mindannyi* `all' can occur with either a non-inflected or inflected form of the preverb:

(24d)	a	fiú	mindannyi**unk**ba	**bele/belénk**	szeretett
	the	boy	all-1PL-ILL	PV-3sg	loved
	`The boy fell in love with each of us'				

This example contrasts with (24c) where the inflected preverb is incompatible with a case-governed lexical NP. Though such an example requires further thought it might be argued that the variant of (24d) with an uninflected preverb and co-occuring quantifier represents an instance where the quantifier is simply an argument of the complex predicate. In contrast, the variant with the inflected preverb might be interpreted as an instance where the quantifier is an adjunct which bears some discourse role to the argument represented by the inflection on the preverb. This second paradigm raises numerous issues concerning how to treat double marking which we cannot address here (see E. Kiss (forthcoming) for some related worries with respect to the present treatment of inflectable preverb constructions) .