Paradigms and Periphrastic Expression:
A Study in Realization-based Lexicalism

Farrell Ackerman
University of California,
San Diego

Gregory Stump
University of Kentucky

> "Paradigms are primarily and mainly of single words
> but where short groups of words or phrases (e.g.,
> Latin, and some Greek, perfective passives,) are
> syntactically comparable to single words in the
> corresponding places of a different paradigm they are
> obviously to be included in paradigms themselves."
> *-Robins 1959:124*

Our intention in this paper is to develop an explanatory account of the special characteristics of periphrastic expressions by refining the traditional notion of PARADIGM employed within inferential-realizational approaches to morphology (Anderson 1992, Aronoff 1994, Zwicky 1990, Stump 1993, 2001). Our proposal draws on this notion in order to develop a substantive reconceptualization of the form and function of the lexicon in constraint-based theories of syntax. In particular, we argue that the paradigms defined by a language's morphology sometimes include periphrastic combinations (PERIPHRASES). As we show, this conclusion affords a simple resolution of a number of neglected problems which periphrases pose for theories of syntax; it also necessitates a radical rethinking of the relation between the form and content of syntactic structures. The proposals developed here are intended to be relatively neutral with respect to the choice among alternative lexicalist theories of syntax (including Lexical Functional Grammar, Head-driven Phrase Structure Grammar, and Construction Grammar); for concreteness, however, we articulate the details of our claims in the terminology of LFG .[1]

1   A lexical-incremental conception of periphrasis

LFG, like other lexicalist and non-lexicalist generativist frameworks, has traditionally subscribed to a morpheme-based conception of morphology.[2] In particular, treatments of various lexical phenomena within this theory have generally assumed what Stump 2001 calls a *lexical-incremental* conception of morphology: such treatments are LEXICAL by virtue of the assumption that affixes, like stems, possess their own separate

---

[1] A morphological (and lexical) account of the data examined here represents a refutation of the claim made with respect to Uralic languages in Mitchell 1993:66 that "… a purely lexical account of inflectional morphology is doomed to failure, even if it is capable of explaining the ordering of morphemes: certain aspects of inflectional morphology can only be explained by means of syntactic processes." The specific syntactic process that Mitchell has in mind is a syntactic blocking mechanism that is hypothesized to account for the distribution of functional information on lexical heads. See Niño 1997 for an LFG treatment of periphrastic negative predicates in Finnish.   See also Ackerman & Webelhuth 1998, Blevins 2001, Börjars et. al. 1997, Koenig 1999, Orgun 1996, Riehemann 2000, Spencer 2001, to appear, and Sadler & Spencer 2001 for comparable proposals.

[2] See Sells (this volume) for a realization-based interpretation of OT-LFG, as well as Ackerman and Webelhuth 1998 for speculation concerning the compatibility of OT-LFG with the sorts of assumptions argued for in the present article.

representations in the lexicon, and they are INCREMENTAL in that the grammatical properties of a fully inflected word are associated with it only as an effect of its acquiring the morphological markers bearing those properties.

In LFG, this approach to morphology was originally employed as a means of defining lexical representations for synthetic word forms functioning as syntactic atoms in c(onstituent)-structure. Thus, a lexically listed verb root such as *tickle* combines with the lexically listed past-tense morpheme *-ed* to yield the inflected word *tickled* which composes or, more technically, unifies the meaning of the root with that of the suffix.

The basic unification apparatus, which served well in handling canonical monotonic relations between words in syntax (e.g. the relation of subject-predicate agreement) as well as in the definition of a synthetic word form's properties, was extended to account for the composition of morphosyntactic property sets expressed by periphrastic combinations of word forms. In proposals such as those of K.P. Mohanan (1982) (for Malayalam auxiliaries) and Falk (1984) (for English *do* support), it was hypothesized that morphosyntactic information from a main verb's lexical representation may join with the information from an auxiliary's lexical representation in the projection of a single f(unctional)-structure representation.[3] Mohanan (1982), for example, postulates the c-structure patterns in (1), in which both a predicative element and an auxiliary are annotated with the equations ↑=↓ (which in LFG identifies these elements as heads).

(1) Malayalam clause structure (Mohanan 1982)

$$S \rightarrow \begin{array}{c} V' \\ \uparrow = \downarrow \end{array}$$

$$V' \rightarrow \begin{array}{cc} X' & V \\ \uparrow = \downarrow & \uparrow = \downarrow \end{array}$$

The effect of this hypothesis is that the lexical information associated with the individual words occupying terminal nodes in the V′ combines to determine the information associated with the f-structure of the V′, hence that associated with the f-structure of the entire sentence. This approach to periphrasis in effect transports the assumptions of lexical-incremental morphology into the domain of syntax: while the morphosyntactic information contained in a predicate's f-structure representation is, in the case of morphologically synthetic forms, projected from a single word's lexical representation via c-structure, it is, in instances of periphrasis, an amalgam of morphosyntactic information distributed among two or more syntactically atomic co-heads in c-structure.

Consider how this lexical-incremental approach to periphrasis might apply in the analysis of inflected predicates in Udmurt. As in other Uralic languages, Udmurt clausal negation is expressed periphrastically by means of a negative verb inflected for subject agreement[4] and a dependent 'main' verb, as in (2). In contrast, the affirmative counterpart of (2) is expressed synthetically, as in (3):

---

[3] See Butt et. al. 1996 for a more recent variant of this hypothesis.
[4] We will see below that there is informative syncretism among negative verb forms.

(2)　　*Ton*　　　　*ud*　　　　　*miniski*　　　(Kel'makov & Hännikäinen 1999:28)
　　　　you　　　　not.2SG　　　go
　　　　'You are not going'

(3)　　*Ton*　　*miniskod*　　　　　　　　　(Kel'makov & Hännikäinen 1999:27)
　　　　you　　go.2SG
　　　　`You are going'

An f-structure representation of the morphologically synthetic predicate *miniskod* in (3) is projected solely from this word form's lexical representation via its presence in c-structure; by contrast, the f-structure representation of the periphrastic predicate *ud miniski* in (2) is, in the lexical-incremental approach to periphrasis, an amalgam of the lexical representations of two distinct syntactic atoms--that of *ud* (4a) and that of *miniski* (4b).[5]
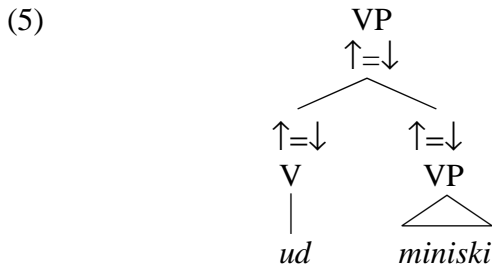
(4)　　a.　　Lexical representation of *ud*
　　　　　　ud　　　　　V

$$
\begin{bmatrix}
\text{POL} & \text{neg} \\
\text{TNS} & \text{pres} \\
\\
\text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{pro} \\ \text{PER} & 2 \\ \text{NUM} & \text{sg} \end{bmatrix}
\end{bmatrix}
$$

　　　　b.　　Lexical representation of *miniski*
　　　　　　miniski　　　V

$$
\begin{bmatrix} \text{PRED} & \text{`go} < \text{SUBJ} >' \end{bmatrix}
$$

Accordingly, the c-structure representation in (5) yields the unified f-structure representation in (6) for the periphrastic construction in (2):

(5)
```
                    VP
                    ↑=↓
                   /    \
           ↑=↓        ↑=↓
            V           VP
            |          /\
           ud      miniski
```

---

(6)

$$\begin{bmatrix} \text{PRED} & \text{`go} < \text{SUBJ} >\text{'} \\ \text{POL} & \text{neg} \\ \text{TNS} & \text{pres} \\ \\ \text{SUBJ} & \begin{bmatrix} \text{PRED pro} \\ \text{PER} & 2 \\ \text{NUM} & \text{sg} \end{bmatrix} \end{bmatrix}$$
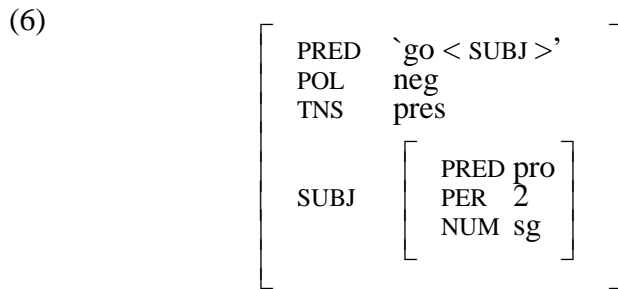
Crucially, in a syntactic treatment of periphrasis such as this there is no single lexical or morphological representation which is associated with the information contained in (6). Instead, an f-structure receives this information as contributions from separate and co-occurring syntactic atoms.

2   An inferential-realizational conception of periphrasis

The supposition that similar sorts of information are supplied to f-structures by single complex lexical representations, as in (3), or multiple lexical entities co-occuring in c-structure, as in (2), has led to the claim within LFG that "morphology competes with syntax." In this paper we argue that for many instances of periphrasis this slogan is not sufficiently discriminating and should be replaced with the claim that in a language's morphology, synthesis may compete with periphrasis as a mode of inflectional exponence--in other words, that periphrases are sometimes actually morphological. The essential ingredients of this proposal are the following two claims:

- A lexeme may be realized synthetically (as a single syntactic atom) or periphrastically (by two or more syntactic atoms co-occurring in a c-structure).
- The contentive information associated with a periphrase is not determined by the contentive information associated with its individual, syntactically independent parts through the mediation of unification principles defined on syntactic structures; rather, the contentive information associated with a periphrase is specified morpholexically. That is, syntactic principles of constituency and linearity determine the distribution of a periphrase's individual parts, but not the functional information which that periphrase expresses.

The antecedents of this competing hypothesis can be found in Ackerman (1984, 1987) (where it is employed in the analysis of Hungarian auxiliary constructions) as well as in the descriptive and pedagogical traditions of numerous languages. This proposal shares with those of Mohanan and Falk (and the recent LFG co-head proposals) the claim that periphrases are associated with a single, simple f-structure (rather than with nested f-structures reflecting the hierarchical organization of a periphrase's c-structure); but it differs dramatically in its assumption concerning the source of the information in that f-structure:

Constructions of these sorts will be referred to as *analytic predicates…* they will be interpreted as entities created by morpholexical rules[6]… a portion of the analytic predicate, specifically the AUX element, functions as a constituent structure verb, i.e., it is the *categorial* or *structural head* of the lexical composition, while the infinitival … contributes the lexical meaning of the derived grammatical word[7], i.e., it serves as the *functional head…*In other words, morpholexical rules produce a grammatical word[8] with discrete structural and functional heads and this analytic composition is associated with a lexical form. (Ackerman 1987:329)

The assumption, in short, was that the clausal predicate information in f-structures is projected from a single lexical representation which receives either synthetic or periphrastic exponence in c-structure.[9]

This approach assumed, but did not develop, a theory of morphology of the type designated as *inferential-realizational* in Stump's taxonomy. Recently there has been a resurgence of interest in the so-called Word & Paradigm approach to morphology (Robins 1959, Matthews 1972, Zwicky 1985, Anderson 1992, Aronoff 1994, Beard 1995, Stump 2001, etc.); what distinguishes this approach from traditional morpheme-based approaches is its premise that a language's inflectional system is INFERENTIAL rather than lexical (in the sense that it represents inflectional exponents not as lexically listed elements, but as markings licensed by rules by which complex word forms are deduced from simpler roots and stems) and is REALIZATIONAL rather than incremental (in the sense that it treats a word's association with a particular set of morphosyntactic properties as a precondition for--not a consequence of--the application of the rules licensing the inflectional exponents of those properties). Inferential-realizational approaches to morphology are in fact quite consistent with the fundamental assumptions of constraint-based lexicalism, both with respect to general conceptual, design features and in their commitment to comprehensive and rigorous formalization of analyses.

The assumption that a periphrase's f-structure is projected from a single lexical representation is not obviously compatible with usual assumptions in LFG concerning the source of information in f-structures. The standard view is, again, that information associated with lexical representations is projected into c-structure and thence, via correspondence algorithms, to f-structures: the f-structure, thus, represents a distillation of syntactically relevant information provided by the lexical elements occupying c-structure terminal nodes. Our claim here, however, is that the syntactic atoms constituting a periphrase may be nothing more than form-theoretic exponents of a unitary content-theoretic element, and that it is this latter element--not its exponents--that

---

[6] Such rules were hypothesized to be relegated to the lexicon and responsible for word-formation and inflection.

[7] See discussion below for the meaning of the term *grammatical word.*

[8] The notion "grammatical word" was intended to reflect Matthew's (1991) hypothesis that wordhood is not a unitary phenomenon and that one dimension of this phenomenon is a contentive one in which the lexical semantics and morphosyntactic properties jointly yield a unit called the "grammatical word." See also Carstairs-McCarthy 2000:596 for a somewhat different view.

[9] Minimally, within LFG the possibility of multi-word lexical items requires modifying the conventions used for annotating c-structure expressions associated with single word lexical items so that appropriate lexical information will produce well-formed f-structures. We leave these sorts of implementational issues to another forum in favor of developing general arguments for the morphological status of periphrases.

determines the periphrase's f-structure. In particular, we claim that rules of morphology define the (potentially periphrastic) realization of a lexeme's pairing with a particular set of morphosyntactic properties, and that the association of such a pairing with an f-structure is insensitive to the manner of its realization.

We develop this inferential-realizational conception of periphrasis in the following sections. In section 3, we argue for a theoretical distinction between two types of paradigms and we discuss the relevance of this distinction to an inferential-realizational theory of periphrasis. In section 4, we examine a range of empirical evidence favoring the proposed theory over syntactically-oriented lexicalist alternatives. Section 4 focuses primarily on the morphosyntactic paradigms of Udmurt and Mari and provides analyses of these predicates in order to demonstrate that the relation between the lexicon and the c-structure is mediated via lexical representations organized in terms of paradigms.

3   Two types of paradigms and the linkage between them

In distinguishing a lexeme's content-theoretic aspects from its form-theoretic aspects, we will pursue an innovative conception of the lexicon and its relation to c-structure, f-structure, and morphological realization. On this conception, a language's lexicon is bipartite with respect to content and form: one part of its lexicon is its LEXEMICON, whose individual entries are lexemes bearing lexical meanings; the complementary part is its RADICON, whose individual entries are roots, i.e., formal elements. Every member L of a language's lexemicon has an associated SYNTACTIC PARADIGM SP(L) such that each cell in SP(L) consists of the pairing of L with a complete set of morphosyntactic properties.[10] Crucially, the cells of these paradigms represent ensembles of semantically interpretable information. In contrast, every member r of a language's radicon has an associated MORPHOLOGICAL PARADIGM MP(r) such that each cell in MP(r) consists of the pairing of r with a set of differentiating morphosyntactic property labels. This paradigm represents the inventory of basic forms used to express the lexemic and morphosyntactic information found in syntactic paradigms.

Consider, for concreteness, the future-tense realizations of the Udmurt verbal lexeme MÏNÏ `go' in Table 1. In our approach, these forms imply the syntactic paradigm in (7) and the morphological paradigm in (8).

---

[10] The relevant sense of completeness here is that of Stump (2001:42f): a set σ of morphosyntactic properties for a lexeme of some category is COMPLETE iff σ is well-formed and for any morphosyntactic property set τ such that σ is not an extension of τ, the unification of τ and σ is not well-formed. Thus, although {masc, nom, sg} is a complete property set for Latin nominal lexemes, {masc, sg} is not; and because a property set can be well-formed only if it doesn't have contrasting values for any feature (p.40f), property sets such as {masc, nom, sg, pl} likewise fail to qualify as complete.

TABLE 1. Future-tense forms of Udmurt MÏNÏ `go'

		Affirmative	Negative
Singular	1	*mïno*	*ug mïnï*
	2	*mïnod*	*ud mïnï*
	3	*mïnoz*	*uz mïnï*
Plural	1	*mïnom(ï)*	*um mïne(le)*
	2	*mïnodï*	*ud mïne(le)*
	3	*mïnozï*	*uz mïne(le)*

(7)　Syntactic future-tense paradigm
　　　of the Udmurt lexeme MÏNÏ `go'
　　a. ⟨ MÏNÏ, {1st singular future affirmative}⟩
　　b. ⟨ MÏNÏ, {2nd singular future affirmative}⟩
　　c. ⟨ MÏNÏ, {3rd singular future affirmative}⟩
　　d. ⟨ MÏNÏ, {1st plural future affirmative}⟩
　　e. ⟨ MÏNÏ, {2nd plural future affirmative}⟩
　　f. ⟨ MÏNÏ, {3rd plural future affirmative}⟩
　　g. ⟨ MÏNÏ, {1st singular future negative}⟩
　　h. ⟨ MÏNÏ, {2nd singular future negative}⟩
　　i. ⟨ MÏNÏ, {3rd singular future negative}⟩
　　j. ⟨ MÏNÏ, {1st plural future negative}⟩
　　k. ⟨ MÏNÏ, {2nd plural future negative}⟩
　　l. ⟨ MÏNÏ, {3rd plural future negative}⟩

(8)　Morphological future-tense paradigm
　　　of the Udmurt root *mïnï* `go'
　　a. ⟨ *mïnï*, {1st singular future affirmative}⟩
　　b. ⟨*mïnï*, {2nd singular future affirmative}⟩
　　c. ⟨*mïnï*, {3rd singular future affirmative}⟩
　　d. ⟨*mïnï*, {1st plural future affirmative}⟩
　　e. ⟨*mïnï*, {2nd plural future affirmative}⟩
　　f. ⟨*mïnï*, {3rd plural future affirmative}⟩
　　g. ⟨ *mïnï*, {1st singular future negative}⟩
　　h. ⟨ *mïnï*, {2nd singular future negative}⟩
　　i. ⟨ *mïnï*, {3rd singular future negative}⟩
　　j. ⟨*mïnï*, {1st plural future negative}⟩
　　k. ⟨*mïnï*, {2nd plural future negative}⟩
　　l. ⟨*mïnï*, {3rd plural future negative}⟩

For each cell ⟨L,σ⟩ in the syntactic paradigm of a lexeme L, there is a corresponding cell in some root's morphological paradigm; we refer to this as the MORPHOLOGICAL CORRESPONDENT (MC) of ⟨L,σ⟩. Thus, ⟨ *mïnï*, {1st singular future affirmative}⟩ in (8) is the MC of ⟨ MÏNÏ, {1st singular future affirmative}⟩ in (7).

The REALIZATION of some cell ⟨r,σ⟩ in the morphological paradigm of a root r is the form defined by the application to r of all relevant morphological rules realizing the property set σ; the realization of some cell ⟨L,σ⟩ in the syntactic paradigm of a lexeme L is the realization of the MC of ⟨L,σ⟩. Thus, the realization of both the cell ⟨ *mïnï*, {1st singular future affirmative}⟩ in (8) and the cell ⟨ MÏNÏ, {1st singular future affirmative}⟩ in (7) is the synthetic word form *mïno* (cf. Table 1); similarly, the realization of both the cell ⟨ *mïnï*, {1st singular future negative}⟩ and the cell ⟨ MÏNÏ, {1st singular future negative}⟩ is the periphrase *ug mïnï*.[11] We do not assume that the MÏNÏ entry in Udmurt's lexemicon necessarily includes the entire syntactic paradigm in (7), nor that the *mïnï* entry in Udmurt's radicon necessarily includes the entire morphological paradigm in (8); we instead assume that these entries simply contain the information necessary for the projection of these paradigms (and of the realizations of their cells) by well-motivated rules of morphology. Thus, although we shall refer to cells such as ⟨ MÏNÏ, {1st singular

---

[11] Periphrastic negative predicates such as these exemplify the construct "expanded predicate" proposed in Ackerman and Webelhuth 1998 (see also Spencer to appear.)

future affirmative}⟩ and ⟨ *mïnï*, {1[st] singular future affirmative}⟩ as aspects of lexical representation, they might be more accurately characterized as "morpholexical".

 The relation between the cells in a root's morphological paradigm and the realizations of those cells is mediated by realizational rules of the sort familiar in inferential-realizational ("Word & Paradigm") theories of morphology. Crucially, however, we also adopt a less usual hypothesis, according to which overt inflectional exponents are not limited to synthetic morphological markings. That is, we adopt the Periphrastic Realization Hypothesis in (9).[12]

(9) The Periphrastic Realization Hypothesis:
 Inflectional rules that deduce the realizations of a morphological paradigm's cells include rules defining periphrastic combinations as well as rules defining synthetic forms.

 While inflectional rules determine the relation between the cells of a root's morphological paradigm and their realizations, the relation between the cells in a lexeme's syntactic paradigm and their morphological correspondents is regulated by a different kind of rule: for any cell ⟨L,σ⟩ in the syntactic paradigm of a lexeme L, the MC of ⟨L,σ⟩ is determined by a RULE OF PARADIGM LINKAGE. In most instances, the operative rule of paradigm linkage is the universal default in (10):

(10) Universal default rule of paradigm linkage:
 If root r is stipulated as the primary root of a given lexeme L, then the MC of ⟨L,σ⟩ is ⟨r,σ⟩.[13]

Thus, if *mïnï* is stipulated as the primary root of the Udmurt lexeme MÏNÏ `go', then by default, the MC of ⟨ MÏNÏ, {1[st] singular future affirmative}⟩ is ⟨ *mïnï*, {1[st] singular future affirmative}⟩ (realization: *mïno*), and so on.

---

[12] See Robins 1959, Matthews 1991, and Haspelmath 2000. See also Börjars et. al. 1997, Sadler & Spencer 2000, and Spencer 2001a,b on the periphrasis of the perfect passive in Latin as well as in various Slavic languages; Stump 2001 on the Sanskrit periphrastic future; and Brassil on Italian (this volume). Though our discussion here focuses on inflectional instances of periphrasis, it should be noted that periphrastic realization is also "extremely common" (Csúcs 1998:295) in the domain of derivation; for instance, Udmurt possesses verbal derivatives consisting of a gerundial form of the main verb and a co-occuring inflected form of a verb encoding properties of tense, aspect, modality, etc. This suggests that synthetic and periphrastic realization obtains for both inflection and derivation and therefore accords with the claims of the Strong Lexicalist Hypopthesis.

[13] Here, we assume that a language's syntactic and morphological paradigms involve the same inventory of morphosyntactic properties. Sadler & Spencer (2001), however, argue for a principled distinction between s(yntactic)-features ("functional features which have to be expressed by well-formed phrases and clauses") and m(orphological)-features ("those that regulate the morphophonological structure of words"); in the context of this distinction, one might assume that each cell ⟨r,σ_m⟩ in a root r's morphological paradigm contains a complete set σ_m of m-feature specifications, but that each cell ⟨L,σ_s⟩ in a lexeme L's syntactic paradigm involves a complete set σ_s of s-feature specifications. On that assumption, the default correspondence between syntactic paradigms and morphological paradigms would have to incorporate the default correspondence of m-feature specifications and s-feature specifications envisioned by Sadler & Spencer (2001:84). This is a possibility we do not exclude; our present concerns do not necessitate this, however.

The default rule of paradigm linkage in (10) can, however, be overridden. Heteroclisis provides a good example. By definition, a heteroclite lexeme is one whose syntactic paradigm contains two distinct cells ⟨L,σ⟩, ⟨L,σ′⟩ such that ⟨L,σ⟩ has the MC ⟨r,σ⟩, ⟨L,σ′⟩ has the MC ⟨r′,σ′⟩, and r ≠ r′. Thus, Sanskrit has a special rule of paradigm linkage which assigns each direct-case (i.e. nominative, vocative, or accusative) cell in the syntactic paradigm of the heteroclite noun HṚD 'heart' a MC from the morphological paradigm of the root *hṛdaya*, while each oblique-case cell in this syntactic paradigm is assigned a MC from the morphological paradigm of the distinct root *hṛd*; the examples in Table 2 illustrate.

TABLE 2. Inflection of the Sanskrit heteroclite noun HṚD 'heart'

Cells in HṚD 's syntactic paradigm	Morphological counterparts	Realizations
⟨ HṚD, {neut <u>nom</u> sg}⟩ ⟨ HṚD, {neut <u>loc</u> sg}⟩	⟨ *hṛdaya*, {neut <u>nom</u> sg}⟩ ⟨ *hṛd*, {neut <u>loc</u> sg}⟩	*hṛdaya-m* *hṛd-i*

Deponence is a phenomenon involving a different sort of override of the default rule of paradigm linkage in (10). A deponent lexeme is, by definition, one whose syntactic paradigm contains a cell ⟨L,σ⟩ whose MC is ⟨r,σ′⟩, where σ ≠ σ′. Thus, Latin has a special rule of paradigm linkage which assigns each active cell in the syntactic paradigm of the deponent verb FATĒRĪ 'confess' a MC from the passive morphological paradigm of FATĒRĪ's primary root *fat*; by contrast, the default rule of paradigm linkage associates cells from the passive morphological paradigm of the primary root *mon* with passive cells in the syntactic paradigm of the nondeponent verb MONĒRE `advise'. The examples in Table 3 illustrate.

TABLE 3. Inflection of the Latin verbs FATĒRĪ `confess' (deponent) and MONĒRE `advise' (nondeponent)

Cells in syntactic paradigms	Morphological counterparts	Realizations
⟨ FATĒRĪ,{1 sg pres <u>act</u> indic}⟩ ⟨ MONĒRE,{1 sg pres <u>pass</u> indic}⟩	⟨ *fat*,{1 sg pres <u>pass</u> indic}⟩ ⟨ *mon*,{1 sg pres <u>pass</u> indic}⟩	*fateor* *moneor*

(See Stump (2002) for additional discussion and exemplification of an inferential-realizational approach to heteroclisis and deponence.)

In the grammatical framework advocated here, the interface between syntactic paradigms and c-structures is regulated by the two principles in (11).

(11) a. Synthetic Realization Principle (= Morphological Expression of Ackerman & Webelhuth 1998):
Where the realization *w* of ⟨L,σ⟩ is a synthetic member of category X, *w* may be inserted as the head of XP.

b. Periphrastic Realization Principle:
Where the realization of w_1 w_2 of $\langle L,\bullet\rangle$ is periphrastic and w_1 and w_2 belong to the respective categories X and Y, w_1 and w_2 may be inserted as the heads of the respective phrases XP and YP.

As formulated, the Periphrastic Realization Principle makes no claims about the surface constituency relations among the elements of periphrastic constructions; thus, our initial assumption is that the structural relationship between XP and YP in (11b) is determined by RULES OF PERIPHRASTIC SYNTAX at least some of which may be language-specific.

According to (11a), the synthetic form *miniskod* `go.2SG' in the Udmurt sentence (3) may head a VP which thereby bears the morphosyntactic property set {2nd singular present affirmative}; according to (11b), the parts of the periphrase *ud miniski* `not.2SG go' in sentence (2) may head separate phrases--presumably the phrases VP_1 and VP_2 in a structure of the form $[_{VP1}$ V $VP_2]$, as in (5). In this way, the individual word forms constituting a realizational periphrase exhibit morphophonological integrity (in accordance with the arguments for Lexical Integrity of Bresnan and Mchombo 1995), even though the realization as a whole fails to satisfy the Synthetic Realization Principle (11a). Thus the Synthetic Realization Principle is interpretable as a violable principle permitting realizations exhibiting variable degrees of analyticity. By the same token, the lexicality of an entity cannot, on our view, be reliably determined by its surface exponence.

The f-structure corresponding to a c-structure defined by (11a,b) is not projected directly from the forms occupying the individual leaves within this c-structure; rather, its definition depends on accessing the information from the syntactic paradigm of a lexeme L--specifically, from the cell $\langle L,\sigma\rangle$ that is realized by the heads in this c-structure. We assume that if $\langle L,\sigma\rangle$ is realized as the periphrase w_1 w_2 in c-structure, then all of the syntactic complement requirements of this periphrase are determined by the lexeme which it realizes (together with the relevant rules of periphrastic syntax). We accordingly conclude that if $\langle L,\sigma\rangle$ has a (synthetic or periphrastic) realization at c-structure and denotes a predicate P, then the subject and complements of $\langle L,\sigma\rangle$'s realization denote P's arguments.

In this framework, the skeletal information associated with a lexeme's f-structural representation is projected strictly from the information in its syntactic paradigm; on the other hand, the c-structural representation of any cell $\langle L,\sigma\rangle$ in the syntactic paradigm of a lexeme L is additionally sensitive to this cell's realization. *Thus, the realizations associated with a syntactic paradigm's (content-theoretic) cells do not themselves participate in the determination of f-structure, but are simply the c-structural expressions of information contained in the cells of syntactic paradigms.* This permits lexical representations to project their information into clauses, independently of how these lexical representations are formally realized. Thus periphrasis can be construed as a purely formal aspect of lexical representations where the pieces themselves are not annotated with information as they are in co-head analyses within LFG. This system can be viewed as an implementation of Realization-based Lexicalism as discussed in Blevins 2001.[14] The basic organization of lexical representations can be represented as in Table 4.

---

[14] In some measure, the proposal espoused here represents an elaboration of the intuitions guiding the Sadler & Spencer 2001 treatment of periphrastic constructions in terms of a distinction between s-structures

TABLE 4. Lexical representations in the assumed framework

	Lexical representations:		Realizations (word forms)
	Cells in syntactic paradigms (<L,σ> pairings)	Cells in morphological paradigms (<r,σ> pairings)	
Associations:	Cells associated with cells by *rules of paradigm linkage*		
		Cells associated with realizations by *realization rules*	
Information represented:	Contentive (functional-semantic content, grammatical functions, morphosyntactic properties)	Formal (morphosyntactic property labels) Diacritic (indexical) (inflection class membership, root phonology)	Purely phonological

4   The lexicality of Mari and Udmurt predicates

In this section we examine instances of periphrasis in two Uralic languages, Mari and Udmurt.   We argue that the evidence from these languages strongly favors a morphological approach to periphrasis--that if principles that are standardly assumed to regulate synthetic exponence in realizational morphological theories are regarded as having periphrastic exponence within their compass as well, then a number of otherwise problematic characteristics of periphrasis turn out to be neither surprising nor unexpected. We develop detailed formal analyses of this data as concrete exemplifications of this approach. Throughout, we explicitly contrast the morphological approach to periphrasis with the `purely syntactic' approach, in which morphological synthesis and periphrasis are theoretically segregated (the former being treated as the domain of a language's morphology and the latter as a province of its syntax).

A priori, the development of a morphological theory of periphrasis confronts a number of important issues. Foremost among these is the issue of criteria: how does one decide whether a given multi-word expression is the realization of some cell in a paradigm? That is, how is a periphrase distinguished from a group of words in a relation of syntactic complementation? Though this issue is virtually ignored in modern linguistic theory, it was central to much speculation on the nature of wordhood in the Soviet linguistic tradition. The following remarks from M. M. Gukhman (1963:199) provide a cogent statement of the basic task:[15]

---

and m-structures: roughly, their s-structure can be construed as the contentive side of the lexical representation embodied by a syntactic paradigm, while their m-structure is expressed by our notion of morphological paradigms (but see footnote 12 above).

[15] See also Zhirmunskij (1963:24).

> The need to establish criteria for the differentiation of analytic verbal constructions from other types of word combinations, such as the combination of two or several full words or the combination of an auxiliary with a full word, is connected with the question of whether these constructions are considered members of a paradigmatic series, that is, whether they are units of the morphological level.

This is a particularly crucial issue, since constraint-based lexicalist frameworks such as LFG and HPSG conventionally assume that periphrasis should not be treated in the lexical/morphological component of grammar, given that it involves syntactically independent elements. As a consequence these frameworks have developed analyses of such phenomena which employ various modifications of the syntactic apparatus used to address other syntactic phenomena. In effect, the default assumption has been that periphrasis falls within the purview of syntactic frameworks since they possess tools that can be modified to treat them. It hardly needs to be said that this argument from parsimony, shared by Chomskyan syntactocentric approaches, represents an analytic convenience rather than an independently motivated, empirically supported, and well-reasoned theoretical proposal. The irony of this position for lexicalist approaches is that, given their adoption of representational modularity, they are not committed architecturally and conceptually to a syntactic analysis of periphrasis in the same way that Chomskyan approaches are.

In the following discussion, we argue that there are at least three sufficient criteria for the identification of periphrases: the criterion of featurally intersective distribution, that of noncompositionality, and that of distributed exponence.[16] Our discussion will therefore focus on analytic combinations which satisfy one or more of these criteria, and our objective, again, is to demonstrate that combinations which satisfy any of these criteria are most plausibly treated as expressions of a language's morphology. Because the languages Mari and Udmurt possess rich inventories of predicates which receive synthetic and periphrastic expression, they are instructive sources of data for the development of this discussion.

Our discussion proceeds as follows. First, we introduce the criterion of featurally intersective distribution and focus attention on a class of analytic combinations from Western Mari which satisfy this criterion (section 4.1); we argue that the purely syntactic approach to periphrasis does not afford an adequate account of the manner in which periphrasis and synthesis are paradigmatically opposed in the Western Mari data; the morphological approach, by contrast, does. We present a detailed formal analysis of a fragment of Western Mari verb morphology to give concrete substance to this claim (section 4.2). In section 4.3, we introduce the criterion of noncompositionality and adduce examples of analytic constructions from Mari and Udmurt which satisfy this criterion; we demonstrate that periphrases of this sort constitute an especially forceful type of evidence against the purely syntactic approach to periphrasis and in favor of the

---

[16] See Spencer 2001 and to appear for the postulation of additional criteria which are useful for distinguishing between morphological versus syntactic objects. Ackerman and Webelhuth 1998, additionally, suggested a general criterion derived from the Strong Lexicalist Hypothesis, namely, that all evidence of the derivational modification of information stands as a sufficient condition of lexicality. This was referred to as The Principle of Lexical Adicity. This principle is consistent with the new, and we believe, more compelling criteria presently being adduced from morphology proper.

morphological approach. In section 4.4, we discuss the criterion of distributed exponence, which we exemplify with additional evidence from Udmurt; as we show, the morphological approach to periphrasis affords a superior account of this evidence as well. We conclude (section 4.5) with a brief discussion of two kinds of evidence which support the morphological approach to periphrasis: the close relation between synthesis and periphrasis in language change and the status of periphrasis in a theory of morphological markedness. Throughout this presentation it should be kept in mind that a morphological approach, while certainly appropriate for some periphrases, may be inappropriate for others. The behaviors considered here may therefore be seen as criteria which either justify or militate against a morphological approach in any given case.

4.1 The paradigmatic opposition of periphrasis to synthesis

Periphrases commonly stand in paradigmatic opposition to synthetic realizations; that is, they realize contrasting values for the same morphosyntactic features but are otherwise identical in their lexicosemantic content. Mari (known also as Cheremis) exhibits numerous instances of this characteristic of periphrasis, some of which we examine here.

Consider, for example, the present desiderative, first-past, and second-past[17] realizations of the verb KOL `die' in Western Mari; these are given in Tables 5-7.[18] The second-past realizations in Table 7 are uniformly synthetic, for both the affirmative-polarity and the negative-polarity portions of the paradigm. The desiderative and first-past realizations in Tables 5 and 6, by contrast, are synthetic in the affirmative but periphrastic in the negative; in particular, each of the negative realizations in Tables 5 and 6 involves a finite form of the negative verb AK (the relevant forms of which are isolated in Table 8) and an invariant form of KOL itself.

TABLE 5. Present desiderative realizations of the Mari *em*-conjugation
verb KOL `die' (Western dialect) [Alhoniemi 1985:125,127]

		Affirmative	Negative
SG.	1	*kolə̂-ne-m* `I want to die'	*ə-ne-m kolə̂* `I don't want to die.'
	2	*kolə̂-ne-t*	*ə-ne-t kolə̂*
	3	*kolə̂-ne-žə*	*ə-ne-žə kolə̂*
PL.	1	*kolə̂-ne-nä*	*ə-ne-nä kolə̂*
	2	*kolə̂-ne-δä*	*ə-ne-δä kolə̂*
	3	*kolə̂-ne-štə*	*ə-ne-štə kolep*

---

[17] According to Kangasmaa-Minn (1998:229), "[t]he first past refers especially to states and events which the speaker has personally witnessed, while the second past is more or less a record of what has been or happened without any emphasis on the speaker's attitude towards the truth value of the utterance".
[18] The segmentation of Mari formatives assumed here and throughout follows the analysis of Eastern Mari verbs proposed by Sebeok & Ingemann (1961).

13

TABLE 6. First-past realizations of the Mari *em*-conjugation verb KOL
`die' (Western dialect) [Alhoniemi 1985:113f, 118]

		Affirmative	Negative
SG.	1	*kolə̂-š-ə̂m* `I died.'	*š-əm kolə̂* `I didn't die.'
	2	*kolə̂-š-ə̂c*	*š-əc kolə̂*
	3	*kolə̂-š*	*əš kolə̂*
PL.	1	*kolə̂-š-na*	*šə-nä kolə̂*
	2	*kolə̂-š-ta*	*šə-δä kolə̂*
	3	*kol-eβ-ə*	*əš kolep*

TABLE 7. Second-past realizations of the Mari *em*-conjugation verb KOL
`die' (Western dialect) [Alhoniemi 1985:114, 118]

		Affirmative	Negative
SG.	1	*kol-en-äm* `I went'	*kolə̂-δe-l-am* `I didn't go'
	2	*kol-en-ät*	*kolə̂-δe-l-at*
	3	*kol-en*	*kolə̂-δe*
PL.	1	*kol-en-nä*	*kolə̂-δe-l-na*
	2	*kol-en-dä*	*kolə̂-δe-l-δa*
	3	*kol-en-ət*	*kolə̂-δe-l-ə̂t*

TABLE 8. Present desiderative and first-past realizations of the Mari
negative auxiliary AK (Western dialect) [Alhoniemi 1985:118]

		Present desiderative	First past
SG.	1	*ə-ne-m*	*š-əm*
	2	*ə-ne-t*	*š-əc*
	3	*ə-ne-žə*	*əš*
PL.	1	*ə-ne-nä*	*šə-nä*
	2	*ə-ne-δä*	*šə-δä*
	3	*ə-ne-štə*	*əš*

The distribution of periphrasis in Tables 5-7 can be construed as FEATURALLY INTERSECTIVE, in the sense that there is no one morphosyntactic property among those expressed by the realizations in Tables 5-7 that is always expressed periphrastically rather than synthetically: not all negative realizations are periphrastic, nor are all desiderative or first-past realizations synthetic; instead, it is the intersection of negative polarity with the desiderative mood or first-past tense which is expressed periphrastically in Western Mari. In particular, since each feature sometimes receives a synthetic realization, it is clear on standard assumptions that they subsume morphosyntactic properties of lexemes and, consequently, that they are properly regarded as within the scope of morphology. Since various combinations of these morphosyntactic properties are associated with

periphrastic exponence, it follows that such phenomena are likewise reasonably regarded as morphological. Featurally intersective distribution is the basis for our first criterion for distinguishing morphologically defined periphrases from other analytic combinations; this criterion is stated in (12). (Note that (12) is merely a sufficient criterion; it does not require that every periphrase have a featurally intersective distribution.)

(12)   Criterion I: If an analytic combination C has a featurally intersective distribution, then C is a periphrase.

By this criterion, all of the analytic combinations in Tables 5 and 6 are periphrases.

       The paradigmatic opposition of synthesis and periphrasis exemplified in Tables 5-7 raises two crucial questions for a theory of periphrasis. First, why do some morphosyntactic property sets lack single-word realizations? And second, why do single-word realizations exclude the use of synonymous periphrases? The purely syntactic approach to periphrasis implies one set of answers to these questions; the morphological approach to periphrasis affords a very different answer. We now show that of the two approaches, the latter provides a superior account of paradigmatic oppositions of synthesis and periphrasis.

4.1.1 Accounting for paradigmatic oppositions in a purely syntactic approach to periphrasis

Consider first the question of why certain morphosyntactic property sets lack single-word realizations. In inferential-realizational approaches to morphology, it is assumed that a language's morphology provides expression for every association of a lexeme's root with a set of morphosyntactic properties available to that lexeme. The expression of such an association may involve overt morphology, as in the realization *walk-ed* expressing the association of the verbal lexeme WALK with the property set {TNS:past, MOOD:indicative, AGR:{PER:1, NUM:pl}}. But overt morphology needn't be involved. For instance, the association of WALK with the property set {TNS:present, MOOD:indicative, AGR:{PER:1, NUM:pl}} is simply expressed by the uninflected root realization *walk*, because English morphology lacks any rule realizing this property set; instances of this sort can therefore be said to exhibit POVERTY OF EXPONENCE.

       The affirmative forms in Tables 5-7 and the negative forms in Table 7 show that properties of polarity, mood, and tense are available to verbal lexemes in Western Mari; thus, in the absence of any contrary stipulation, property sets combining negative polarity with desiderative mood or first-past tense should be available to verbal lexemes in Western Mari. Yet, as we have seen, there is no overt synthetic expression of the association of a verb root with a property set specified for negative polarity and either desiderative mood or first-past tense, nor are such associations simply expressed (through poverty of exponence) as uninflected root forms. Why is this?

       In a purely syntactic approach to periphrasis in which all synthetic realizations are defined by morphological realization rules and all analytic combinations are defined by ordinary principles of syntax, the absence of one-word realizations in the negative desiderative or negative first past (whether these be synthetically inflected realizations or, through poverty of exponence, simple uninflected forms) would have to be attributed to the following ad hoc stipulation:

(13)   Contrary to expectation, property sets specified for negative polarity and either desiderative mood or first-past tense are not available to Mari verbal lexemes.

On this view, a lexeme's morphological paradigm would only contain cells for those property sets that are realized synthetically. At the same time, a purely syntactic approach would have to ensure that the syntax of Western Mari would define negative desiderative and negative first-past periphrases; as a consequence, the fact that the negative desiderative and negative first past are realized periphrastically and not synthetically would be improbably portrayed as a coincidental effect of morphological restrictiveness and syntactic permissiveness.

The second critical question raised by the paradigmatic opposition between periphrasis and synthesis concerns the fact that a synthetic realization overrides the use of a synonymous periphrase: why, for example, does each instance of synthesis in Tables 5-7 exclude the use of a periphrastic alternative? A priori, there is nothing about the syntactic approach to periphrasis that entails that property sets expressed by periphrastic means should strictly complement those property sets that are realized synthetically. Proponents of this approach must therefore stipulate this complementarity by appealing to an ad hoc principle of "morphological blocking" (Andrews 1990, Blevins 2000). According to such a principle, the existence of a synthetic expression for a specified morphosyntactic property set excludes the use of a synonymous periphrase. This principle reifies the basic bifurcation of synthetic and analytic marking underlying the purely syntactic approach to periphrasis, stipulating an otherwise unanticipated domain of competition between morphology and syntax; the need to appeal to such a principle is thus an artifact of the purely syntactic approach to periphrasis. On the face of it, the introduction of such a principle acknowledges that the availability of certain syntactic expressions is determined by morphology, but indirectly (rather than directly, as in the account we propose): part of a language's syntax must be treated as effectively the complement of morphology that does not receive synthetic exponence.

4.1.2   Accounting for paradigmatic oppositions in a morphological approach to periphrasis

A morphological approach to periphrasis provides a much less stipulative account of the paradigmatic opposition of synthesis and periphrasis. In this approach, morphological rules of synthesis and periphrasis participate competitively--as alternatives--in the realizational definition of a lexeme's forms. The competition, however, is not between morphology and syntax, as in morphological blocking proposals, but between the varieties of exponence employed in realizing the cells in a lexeme's syntactic paradigm. The fact that some morphosyntactic property sets lack single-word realizations is therefore attributed to (a) the lack of any rules of synthetic exponence realizing those property sets and (b) the existence of a general default rule of periphrastic exponence realizing those sets. No ad hoc stipulation comparable to (13) is needed in this approach: property sets specified for negative polarity and either desiderative mood or first past tense are, as expected, available for realization in the inflection of Western Mari verbal

lexemes, but happen to be realized by a default rule of periphrasis.[19]   Thus, the morphological approach is--unlike the syntactic approach--fully compatible with the restrictive hypothesis that any property set that is legal in c-structure can also legally drive morphological realization.   The fact that synthesis excludes periphrasis in the negative second past follows from the assumption that the default rule of periphrasis defining negative realizations is (in accordance with Pāṇini's principle) overridden by a narrower rule of synthesis defining negative second-past realizations; thus, the morphological approach also avoids appealing to an ad hoc principle of morphological blocking to account for the exclusion of periphrasis in such instances.   In other words, there is no need to posit a special blocking principle to regulate the relation between morphology and syntax: as a general, independently motivated constraint within morphology, Pànini's principle suffices to account for the relevant data if periphrasis is defined morphologically.

4.2   Formal analysis of a fragment of Western Mari verb morphology

We now develop a formal analysis of the Mari realizations in Tables 5-7 which embodies these advantages of the morphological approach to periphrasis and which will provide a concrete basis for further discussion. At the core of this analysis is our assumption (noted abover, cf. Table 4) that the cells in a lexeme's morphological paradigm are associated with their realizations by a systems of realization rules.

4.2.1   Basic assumptions

The analysis that we shall propose for the Mari verb forms in Tables 5-7 rests on a number of pretheoretic assumptions, which we now elucidate. Morphologically, the desiderative and first-past realizations in Tables 5 and 6 are built upon KOL's SCHWA STEM, i.e. the result of suffixing schwa to KOL's root *kol*;[20] exceptionally, however, the third-person plural form of the first past is built on a special stem in *-ep* (prevocalic alternant: *-eβ*). The schwa stem *kolə̂* and the special stem *kolep* will be referred to as KOL's PRIMARY STEMS. In the affirmative of the present desiderative and of the first past, the primary stem is augmented by the modal suffix *-ne* and the temporal suffix *-š* (respectively); the *-š* suffix is, however, grammatically restricted to *em*-conjugation verbs and phonologically restricted to postvocalic positions (and is, for this latter reason, absent from the third-person plural affirmative first-past form *kol-eβ-ə*, whose primary stem ends in a consonant). Augmenting a primary stem in these ways produces a SECONDARY STEM. A verb form's subject-agreement terminations are affixed to its secondary stem; in the first past, however, the third person singular is a default realization lacking any overt personal termination. In the periphrastic negative realizations of the present desiderative and first past, the inflections for mood and subject agreement appear not on the stem of KOL itself, but on a primary stem of the negative auxiliary AK (which, though negative in

---

[19] The default appeal to periphrastic realization parallels the status of periphrastic negative expressions as the unmarked encoding within Uralic.
[20] Because other forms of KOL are also built on this stem and because the full set of forms that are built on it are not unified by any common morphosyntactic property, `schwa-stem' is a morphomic category (Aronoff 1994.)

meaning, exhibits the same markers as used in affirmative morphology). The primary stems of AK are irregular: its primary desiderative stem is *ə*; its primary first-past stem is *əš* in third-person realizations, otherwise *šə*.[21]

The affirmative second-past realizations in Table 7 are built on a primary stem identical to KOL's affirmative gerundial stem (which consists of its root plus the *em*-conjugation gerundial suffix *-en*). This primary stem is inflected with person/number markers expressing subject agreement; here also, the third-person singular realization carries no overt agreement marking.[22] The negative second-past realizations are built on a primary stem having two distinct forms--an ABSOLUTE form identical to its negative gerundial stem (its root plus the negative marker *-δe*) and a CONJUNCT form arising from the absolute form through the suffixation of *-l*. The conjunct stem is used in the presence of inflectional suffixes and the absolute form, in their absence; thus, the latter form is restricted to the third-person singular, since it alone is not overtly marked for subject agreement.

Notwithstanding the different degrees of morphological synthesis exhibited by the present desiderative, first-past, and second-past realizations, our assumption is that their syntactic paradigms are parallel; the relevant differences between these tenses are differences in the modes of exponence exploited in the realization of the syntactic paradigm's cells (through the realization of these cells' MCs in the corresponding morphological paradigm).

We assume that the relation between a cell in a morphological paradigm and the realization of that cell is mediated by realization rules; following Stump (2001:44), we assume the following format for realization rules.

(14) Format for realization rules: $RR_{n,\tau,C}(\langle X, \sigma \rangle) =_{def} \langle Y, \sigma \rangle$

A rule $RR_{n,\tau,C}$ stated in this format is to be interpreted as follows: Given a pairing $\langle X, \sigma \rangle$ consisting of a root or stem belonging to class C and a property set σ that is an extension of τ, the result of applying the rule to $\langle X, \sigma \rangle$ is, by definition, the pairing $\langle Y, \sigma \rangle$. **For example, given the entry from the morphological paradigm** ⟨ kol, {pres, desiderative, 1sg}⟩, **where the value for C is verb and where the specified morphosyntactic properties are within the extension of the morphosyntactic properties appropriate for verbs, the realization of this entry, perhaps after the application of several realization rules, will be** kolə̂nem.

The *n* subscript in the rule schematized in (14) identifies it as a member of rule-block *n*. In general, a realization rule belonging to a particular rule-block competes with other members of the same block in the definition of a cell's realization. Rule competition of this sort is resolved by Pāṇini's principle: when two realization compete in the definition

---

[21] There is clearly a historical connection between the sibilant appearing in the first-past stem of AK and the first-past suffix *-š*; but in view of the idiosyncratic variation in the shape of the former, it is not clear that these should be synchronically identified.

[22] Bereczki (1990:55) reports that in some eastern dialects, the first- and second-person plural affirmative second-past forms are instead periphrastic, consisting of an uninflected gerund and a copula inflected both for present tense and for subject agreement, e.g. *tolə̂n ulə̂na* `we came', *tolə̂n ulə̂δa* `you came'.

of a cell's realization, the narrower rule prevails.[23] In such instances, the prevailing rule can be referred to by means of the Nar_n notation (Stump 2001:52): where $\text{RR}_{n,\tau,C}$ is the narrowest rule in block *n* which is applicable to $\langle X, \sigma \rangle$, the notation '$\text{Nar}_n(\langle X, \sigma \rangle)$' represents the result of applying $\text{RR}_{n,\tau,C}$ to $\langle X, \sigma \rangle$. This notation will be useful for defining the systematic resolution of competition among rules of synthetic and periphrastic exponence.

Our assumption is that three rule blocks are necessary for the definition of the Mari realizations in Tables 5-7. The first block (here labelled `Block I') houses rules which deduce a verb's primary stem forms from its root; the second (labelled `Block II') houses rules which allow a verb's secondary stem to be deduced from its primary stem; the third (`Block III') houses the various rules specifying the exponents (if any) of subject agreement. On this assumption, we can say that for any cell $\langle R, \sigma \rangle$ in a Mari morphological paradigm, W is the realization of $\langle R, \sigma \rangle$ if and only if $\text{Nar}_{\text{III}}(\text{Nar}_{\text{II}}(\text{Nar}_{\text{I}}(\langle R, \sigma \rangle))) = \langle W, \sigma \rangle$.[24]

4.2.2 Mari realization rules

In this framework, Western Mari may be seen as having the three blocks of realization rules in (15)

(15) Some realization rules for present desiderative, first-past, and second-past verb forms in Mari (Western dialect) [N.B.: The variable X ranges over stems but not periphrases.]

Block I [Rules deducing primary stems from roots]

W-Ia. $\text{RR}_{\text{I},\{\},\text{V}}(\langle X, \sigma \rangle) = \langle Y, \sigma \rangle$, where Y is X's schwa stem.

W-Ib. $\text{RR}_{\text{I},\{\text{POL:aff, TNS:1st past, AGR:\{PER:3, NUM:pl\}\}},\text{V}}(\langle X, \sigma \rangle) = \langle Xep, \sigma \rangle$.

W-Ic. $\text{RR}_{\text{I},\{\text{TNS:2nd past}\},\text{V}}(\langle X, \sigma \rangle) = \langle Y, \sigma \rangle$, where Y is the realization of $\langle X, \{\text{VFORM:gerund, POL:aff}\}\rangle$.

W-Id. $\text{RR}_{\text{I},\{\text{POL:negative}\},\text{V}}(\langle X, \sigma \rangle) = \langle [\underline{Y}\ Z], \sigma \rangle$, where Y is the realization of $\langle \text{AK}, \sigma \rangle$, $\text{Nar}_{\text{I}}(\langle X, \sigma' \rangle) = \langle Z, \sigma' \rangle$, and $\sigma' = \sigma / \{\text{POL:aff, TNS:1}^{st}\ \text{past}\}$.

W-Ie. $\text{RR}_{\text{I},\{\text{POL:neg, TNS:2nd past}\},\text{V}}(\langle X, \sigma \rangle) = \langle Y, \sigma \rangle$, where Y is an absolute stem form which
        (a) is identical in form to the realization of $\langle X, \{\text{VFORM:gerund, POL:neg}\}\rangle$
  and
        (b) has Y*l* as its conjunct stem form.

W-If. $\text{RR}_{\text{I},\{\text{VFORM:gerund; POL:aff}\},\text{V[CONJUGATION:em]}}(\langle X, \sigma \rangle) = \langle Xen, \sigma \rangle$.

W-Ig. $\text{RR}_{\text{I},\{\text{VFORM:gerund; POL:neg}\},\text{V[CONJUGATION:em]}}(\langle X, \sigma \rangle) = \langle Y\delta e, \sigma \rangle$.

W-Ih. $\text{RR}_{\text{I},\{\text{MOOD:desiderative}\},\{\text{AK}\}}(\langle X, \sigma \rangle) = \langle \partial, \sigma \rangle$.

W-Ii. $\text{RR}_{\text{I},\{\text{TNS:1st past}\},\{\text{AK}\}}(\langle X, \sigma \rangle) = \langle š\partial, \sigma \rangle$.

W-Ij. $\text{RR}_{\text{I},\{\text{TNS:1st past, AGR:\{PER:3\}\}},\{\text{AK}\}}(\langle X, \sigma \rangle) = \langle \partial š, \sigma \rangle$.

---

[23] The relevant notion of narrowness is that of Stump (2001:52): $\text{RR}_{n,\sigma,C}$ is NARROWER than $\text{RR}_{n,\tau,C}$ iff σ is an extension of τ and $\sigma \neq \tau$; where $C \neq C'$, $\text{RR}_{n,\sigma,C}$ is NARROWER than $\text{RR}_{n,\tau,C'}$ iff $C \subseteq C'$.

[24] This generalization constitutes one clause in the definition of Western Mari's paradigm function; cf. Stump (2001:50ff).

Block II [Rules deducing secondary stems from primary stems]

W-IIa. $RR_{II,\{\text{POL:affirmative, TNS:1st past}\},V[\text{CONJUGATION:em}]}(\langle X, \sigma\rangle) = \langle X\check{s}, \sigma\rangle$, where X is vowel-final.

W-IIb. $RR_{I,\{\text{POL:affirmative, MOOD:desiderative, TNS:present}\},V}(\langle X, \sigma\rangle) = \langle Xne, \sigma\rangle$.

Block III [Rules expressing subject agreement]

W-IIIa. $RR_{III,\{\text{POL:affirmative, TNS:1st past, AGR:\{PER:1, NUM:sg\}}\},V}(\langle X, \sigma\rangle) = \langle X\partial m, \sigma\rangle$.

W-IIIb. $RR_{III,\{\text{POL:affirmative, TNS:1st past, AGR:\{PER:2, NUM:sg\}}\},V}(\langle X, \sigma\rangle) = \langle X\partial c, \sigma\rangle$.

W-IIIc. $RR_{III,\{\text{TNS:2nd past, AGR:\{PER:1, NUM:sg\}}\},V}(\langle X, \sigma\rangle) = \langle Xam, \sigma\rangle$.

W-IIId. $RR_{III,\{\text{TNS:2nd past, AGR:\{PER:2, NUM:sg\}}\},V}(\langle X, \sigma\rangle) = \langle Xat, \sigma\rangle$.

W-IIIe. $RR_{III,\{\text{TNS:2nd past, AGR:\{PER:3, NUM:pl\}}\},V}(\langle X, \sigma\rangle) = \langle X\hat{\partial}t, \sigma\rangle$.

W-IIIf. $RR_{III,\{\text{AGR:\{PER:1, NUM:sg\}}\},V}(\langle X, \sigma\rangle) = \langle Xm, \sigma\rangle$.

W-IIIg. $RR_{III,\{\text{AGR:\{PER:2, NUM:sg\}}\},V}(\langle X, \sigma\rangle) = \langle Xt, \sigma\rangle$.

W-IIIh. $RR_{III,\{\text{AGR:\{PER:1, NUM:pl\}}\},V}(\langle X, \sigma\rangle) = \langle Xna, \sigma\rangle$.

W-IIIi. $RR_{III,\{\text{AGR:\{PER:2, NUM:pl\}}\},V}(\langle X, \sigma\rangle) = \langle Xda, \sigma\rangle$.

W-IIIj. $RR_{III,\{\text{AGR:\{PER:3, NUM:pl\}}\},V}(\langle X, \sigma\rangle) = \langle X\partial, \sigma\rangle$.

W-IIIl. $RR_{III,\{\text{POL:affirmative, MOOD:desiderative, AGR:\{PER:3, NUM:sg\}}\},V}(\langle X, \sigma\rangle) = \langle X\check{z}\partial, \sigma\rangle$.

W-IIIm. $RR_{III,\{\text{POL:affirmative, MOOD:desiderative, AGR:\{PER:3, NUM:pl\}}\},V}(\langle X, \sigma\rangle) = \langle X\check{s}t\partial, \sigma\rangle$.

W-IIIn. $RR_{III,\{\text{TNS:1st past, AGR:\{PER:3, NUM:pl\}}\},\{AK\}}(\langle X, \sigma\rangle) = \langle X, \sigma\rangle$.

The rules in Block I (rules **W-Ia** through **W-Ij**) define the primary stems of the realizations in Tables 5-7. By rule **W-Ia**, a verb's schwa stem is its default primary stem. By Pāṇini's principle, however, this default is overridden any time it competes with another rule in Block I. By **W-Ib**, affirmative third-person plural first-past forms have a primary stem in *-ep*, and by the rule of referral[25] **W-Ic**, a verb's second-past forms have a primary stem identical to that verb's affirmative gerund.

Rule **W-Id** is central to our analysis of the periphrasis in Tables 5 and 6. By this rule, the default realization of any negative cell $\langle X,\sigma\rangle$ in a verb's morphological paradigm is a periphrase. Here and throughout, we represent periphrases in brackets: [Y Z]. We further assume that in general, periphrases are headed, and we identify the head of a periphrase by underlining: [Y Z]. The periphrases defined by rule **W-Id** are subject to rule (16) at the morphology-syntax interface; we assume for present purposes that this is a language-specific rule.

(16) Periphrases at the morphology-syntax interface:
  Where [Y Z] (or [Z Y]) is the realization of a cell $\langle R, \sigma\rangle$ in a morphological paradigm such that R belongs to category X, then in c-structure, Y heads XP[σ] and Z heads an XP complement of Y.

The slash notation $\sigma/\{F_1:\rho_1, ..., F_n:\rho_n\}$ in **W-Id** refers to that morphosyntactic property set that is like σ except that its value for the feature F_i is ρ_i $(1 \leq i \leq n)$. (See Stump (2001:56) for a more precise definition.) Thus, the rule of referral **W-Id** causes a negative form to be

---

[25] See Zwicky 1985 and Stump 1993, 2001 for discussion of the special properties of rules of referral.

realized--by default--as the periphrastic combination of a finite form of the negative verb AK with the same primary stem of the corresponding affirmative first-past realization; note that this rule applies whether the primary stem of the corresponding affirmative first-past realization is a schwa stem (defined by rule **W-Ia**) or a stem in *-ep* (defined by rule **W-Ib**). In this way, rule **W-Id** applies to pairings such as (17a,b) in the morphological paradigm of KOL `die' to define the respective realizations (18a,b); by (16), these realizations have the c-structure representations in (19a,b).

(17) a. $\langle kol$, {POL:neg, MOOD:indicative, TNS:1st past, AGR:{PER:1, NUM:pl}}$\rangle$

  b. $\langle kol$, {POL:neg, MOOD:desiderative, TNS:present, AGR:{PER:3, NUM:pl}}$\rangle$

(18) a. [*šə̱nä kolə̂*]

  b. [*ə̱neštə kolep*]

(19) a. [$_{VP}$ [$_V$ *šə̱nä*][$_{VP}$ [$_V$ *kolə̂*]]]

  b. [$_{VP}$ [$_V$ *əneštə*][$_{VP}$ [$_V$ *kolep*]]]

Because the variable X in (15) ranges over stems (including roots) but never periphrases, the form defined by **W-Id** is not itself subject to any rule in (15); instead, it is subject only to the Identity Function Default (20), a universal realization rule acting as the ultimate default in every rule block of every language (Stump 2001:53).

(20) Identity Function Default: Where Y ranges over stems *and* periphrases,
    $RR_{n,\{\},U}(\langle Y, \sigma\rangle) = \langle Y, \sigma\rangle$.

In the inflection of negative second-past forms, rules **W-Ia**, **W-Ib**, and **W-Id** are all overridden by the rule of referral **W-Ie**, according to which the primary stem of a negative second-past verb form has an absolute form Y identical to the verb's negative gerund and a conjunct form resulting from the suffixation of *l* to Y.

Rules **W-Ic** and **WI-e** refer a verb's affirmative and negative gerund forms; these are defined by rule **W-If** and **W-Ig**. By rule **W-If**, an *em*-conjugation verb's affirmative gerund arises from its root through the suffixation of *-en*; by **W-Ig**, its negative gerund arises through the suffixation of *-δe* to its schwa stem.

Rules **W-Ih** through **W-Ij** account for the morphological irregularity of the negative verb AK: **W-Ih** identifies *ə* as AK's primary desiderative stem; **W-Ii** identifies *šə* as the default form of AK's primary first-past stem; and **W-Ij** identifies *əš* as the primary stem for third-person first-past forms of AK. A careful examination of the realizations in Tables 5, 6, and 8 reveals that the sequence of suffixes used to realize tense, mood, and agreement in each affirmative realization is virtually identical to the auxiliary form introducing the corresponding negative realization.[26] Because the forms of the negative

---

[26] This state of affairs is described in the pedagogical grammar of Mari by Zorina, Z. G. et.al. (1990:114) in the following way:

auxiliary are morphologically affirmative, it must be seen as a "deponent" verb (cf. the discussion surrounding the Latin forms in Table 3 in section 3): each cell in its syntactic paradigm contains the property POL:negative, while the corresponding cell in its morphological paradigm instead contains the property POL:affirmative, as in the examples in Table 9.

TABLE 9. Examples of the inflection of the Western Mari negative auxiliary AK (deponent)

Cells in syntactic paradigms	Morphological counterparts	Realizations
$\langle$AK,{<u>POL:neg</u>, TNS:pres, AGR:{PER:1, NUM:sg}}$\rangle$	$\langle ak$,{<u>POL:aff</u>, TNS:pres, AGR:{PER:1, NUM:sg}}$\rangle$	*am*
$\langle$AK,{<u>POL:neg</u>, TNS:1st past, AGR:{PER:2, NUM:pl}}$\rangle$	$\langle ak$,{<u>POL:aff</u>, TNS:1st past, AGR:{PER:2, NUM:pl}}$\rangle$	*šə̆ðä* (cf. Table 8)

Thus, we assume that the relation between AK's syntactic and morphological paradigms is regulated by the rule of paradigm linkage in (21).

(21)   Rule of paradigm linkage for AK:
         Where σ is an extension of {POL:neg}, the MC of $\langle ak$, σ$\rangle$ is $\langle ak$, σ/{POL:aff}$\rangle$.

The Block II rules **W-IIa** and **W-IIb** determine the secondary stems used in a verb's present desiderative, first-past, and second-past realizations. By **W-IIa**, a secondary stem in -*š* is used in defining the affirmative first-past realizations of a verb belonging to the *em*-conjugation; by **W-IIb**, a secondary stem in -*ne* is used in defining a verb's affirmative present desiderative realizations. The Identity Function Default guarantees that for any realization whose secondary stem is not determined by **W-IIa** or **W-IIb**, the relevant secondary stem is simply identical to the primary stem defined by Block I.

The outer layer of a verb's inflectional morphology is regulated by the rules of exponence in Block III: by rule **W-IIIa**, a first-person singular affirmative first-past realization involves the suffixation of -*ə̆m* to the secondary stem defined by Block II; by rule **W-IIIb**, a second-person singular affirmative first-past realization involves the suffixation of -*at*; and so on. Note that rules **W-IIIf** through **W-IIIj** introduce default expressions of agreement; others are additionally restricted according to both polarity and tense (**W-IIIa**, **W-IIIb**), to both polarity and mood (**W-IIIl**, and **W-IIIm**), or to tense alone (**W-IIIc**, **W-IIId**, **W-IIIe**, **W-IIIn**). Because of the Identity Function Default, the absence of any Block III rule realizing third-person singular subject agreement outside of

         "The negative form of the 1st past tense is formed with the help of negative words,
         which are represented by the suffixal part of the 1st past tense of 2nd declension verbs
         and the stem of the imperative."

It should also be noted that this strategy for negative formation is applicable to 1st declension verbs as well, despite the fact that the suffixes in the affirmative 1st past for this class differ from those of 2nd class.

the desiderative entails that a verb's third-person singular first- and second-past realizations will simply lack overt agreement morphology.[27]

4.2.3   Analysis summary

In this analysis of Mari verb morphology, periphrasis (as introduced by rule **W-Id**) is treated as a kind of morphological exponence. Within the morphology itself, periphrasis is theoretically unremarkable, serving alongside various synthetic devices as just another kind of exponence available to inflectional systems.   The distinctive character of periphrasis emerges only at the morphology/syntax interface, where the c-structure representation of a periphrase is regulated by the special rule in (16). The paradigmatic opposition of synthesis and periphrasis follows as a necessary consequence of this analysis. The fact that some morphosyntactic property sets lack single-word realizations is not attributed to an ad hoc gap (comparable to (13)) in the inflection of Mari verbs, but is attributed to (a) the lack of any rules of synthetic exponence realizing those property sets and (b) the existence of a default rule of periphrastic exponence realizing those sets. The fact that single-word realizations exclude the use of synonymous periphrases is not attributed to an ad hoc principle of morphological blocking, but follows, more generally, as a direct consequence of the way in which Pāṇini's principle regulates realizational morphology.

4.3   Noncompositional periphrases

In Soviet linguistics there is a tradition of distinguishing between an analytic word combination belonging to a lexeme's paradigm and a word combination whose parts stand in a purely syntactic relationship; one of the most reliable and compelling diagnostic criteria for distinguishing combinations of the former type is the noncompositionality of the meanings associated with the individual words of which they are constituted.[28] Thus, M. M. Gukhman (1955:343) concludes with respect to German periphrastic verbal constructions:

> "…the grammatical meaning of analytic constructions in German is never equal to the sum of the grammatical meanings of its component parts, but appears as the meaning of an nondecomposable whole."

We regard this noncompositionality as a second sufficient criterion for the identification of periphrases:

---

[27] A sketch concerning some of the morphophonological properties associated with the realization rules for Mari is provided in the Appendix.

[28] It is important to convey that by appealing to this criterion we are not repudiating the notion of semantic compositionality: the hypothesis that words are provided with fully specified feature sets defining their grammatical meanings independent of their realization simply means in general that such meanings are not determined by forms. As mentioned below, it is an important task to develop a semantics of grammatical meaning which does not depend upon morpheme-based assumptions.

(22)   Criterion II: If the morphosyntactic property set associated with an analytic combination C is not the composition of the property sets associated with its parts, then C is a periphrase.

The usual view of such composition is that the content of a complex expression follows from the content of its immediate constituents by a principle of property unification; this is consistent with the general lexical-incremental approach to periphrasis adopted in lexicalist frameworks and elsewhere. In contrast, the Periphrastic Realization Hypothesis implies that a periphrase's membership in a paradigm may have independent consequences for this computation. And indeed, close inspection of periphrastic constructions in Uralic reveals that while some such constructions can be claimed to be compositional in their content, not all can.

    As an instance of apparent compositionality, consider again the negative first-past forms of Western Mari KOL `die' in Table 6; each of these is construable as an analytic combination of a stem of KOL with a first-past form of the negative auxiliary AK. Ordinary principles of property unification appear to guarantee that in c-structure, a negative first-past verb phrase such as [$_{VP}$ *šəδä kolə̂* ] `you didn't die' will be associated with the desired morphosyntactic property set (namely {POL:neg, TNS:1st past, AGR:{PER:2, NUM:pl}}), since this is the very property set associated with the verb phrase's head (the negative auxiliary) in its syntactic paradigm.

    In many other instances, however, ordinary assumptions about property unification do not suffice to determine the content of a periphrase from the content of its parts. Given this, the challenge for theory construction, of course, is to identify the most principled way of accounting for both apparently compositional and non-compositional phenomena. Consider, for instance, the negative second-past forms of KOL `die' in the Eastern dialects of Mari in Table 10: each of these is an analytic combination of KOL's affirmative gerundial stem *kolen* with a negative present-tense form of the copula UL (cf. Table 11); this latter form is itself a compound of a present-tense form of the negative auxiliary OK (cf. Table 12) with the stem *ə̂l* of the copula UL.

TABLE 10. Second-past realizations of the Mari *em*-conjugation verb KOL `die'
(Eastern dialects) [Alhoniemi 1985:110,116]

		Affirmative	Negative
SG.	1	*kol-en-am* `I died'	*kolen omə̂l* `I didn't die'
	2	*kol-en-at*	*kolen otə̂l*
	3	*kol-en*	*kolen oγə̂l*
PL.	1	*kol-en-na*	*kolen onal*
	2	*kol-en-da*	*kolen oδal*
	3	*kol-en-ə̂t*	*kolen oγə̂tə̂l*

TABLE 11. Present-tense realizations of the Mari *am*-conjugation
copula UL[29] (Eastern dialects) [Alhoniemi 1985:111,116]

		Affirmative	Negative
SG.	1	*ul-a-m* `I am'	*om-ə̂l* `I am not'
	2	*ul-a-t*	*ot-ə̂l*
	3	*ul-eš*	*oγ-ə̂l*
PL.	1	*ul-ə̂-na*	*ona-l*
	2	*ul-ə̂-δa*	*oδa-l*
	3	*ul-ə̂-t*	*oγə̂t-ə̂l*

TABLE 12. Present-tense realizations of the Mari negative auxiliary OK[30]
(Eastern dialects) [Alhoniemi 1985:115f]

SG.	1	*o-m*
	2	*o-t*
	3	*ok*
PL.	1	*o-na*
	2	*o-δa*
	3	*oγ-ə̂-t*

Thus, none of the parts of a negative second-past verbal periphrase expresses the second-past tense; indeed, the exponents of tense carried by the finite head of such a periphrase are expressions of the present tense. Consequently, though the verb phrase [VP *kolen omə̂l*] `I didn't die' is associated with the property set {POL:neg, TNS:2nd past, AGR:{PER:1, NUM:sg}}, this association cannot be seen as an effect of ordinary property unification. The analytic combinations in Table 10 are therefore periphrases by criterion (22).

In a morphological approach to periphrasis, the association of the periphrase *kolen omə̂l* `I didn't die' with the property set {POL:neg, TNS:2nd past, AGR:{PER:1, NUM:sg}} is effected morphologically. In particular, we propose that the morphology of Eastern Mari defines the second-past realizations of KOL by means of the realization rules in (23). As in our earlier discussion of Western Mari, we assume that for any cell $\langle R, \sigma \rangle$ in a Mari morphological paradigm, W is the realization of $\langle R, \sigma \rangle$ if and only if $\text{Nar}_{\text{III}}(\text{Nar}_{\text{II}}(\text{Nar}_{\text{I}}(\langle R, \sigma \rangle))) = \langle W, \sigma \rangle$; we further assume that the Eastern Mari negative auxiliary OK is subject to

---

[29] The plural negative realizations in (15) exhibit some variation in shape; see section 4.5.1 below.

[30] The present-tense forms of OK sometimes follow the *am*-conjugation, as in the third-person singular realization *oγ-eš* and the first- and second-person plural realizations as *oγ-ə̂-na* and *oγ-ə̂-δa* (Alhoniemi 1985:115f).

a rule of paradigm linkage analogous to (21), and that OK's default root form *ok* has *o* as its regular preconsonantal alternant.[31]

(23) Some realization rules for second-past verb forms in Mari (Eastern dialects)

<table>
<tr><td></td><td></td><td>(Comparable rule in (15))</td></tr>
<tr><td colspan="2">**Block I**</td><td></td></tr>
<tr><td>**E-Ia.**</td><td>$RR_{I,\{TNS:2nd\ past\},V}(\langle X, \sigma\rangle) = \langle Y, \sigma\rangle$,
where Y is the realization of $\langle X, \{VFORM:gerund; POL:aff\}\rangle$.</td><td>(= **W-Ic**)</td></tr>
<tr><td>**E-Ib.**</td><td>$RR_{I,\{VFORM:gerund;\ POL:aff\},V[CONJUGATION:em]}(\langle X, \sigma\rangle) = \langle Xen, \sigma\rangle$.</td><td>(= **W-If**)</td></tr>
<tr><td>**E-Ic.**[32]</td><td>$RR_{I,\{POL:affirmative,\ TNS:present,\ AGR:\{PER:3,\ NUM:pl\}\},V}(\langle X, \sigma\rangle) = \langle X\hat{\partial}, \sigma\rangle$.</td><td></td></tr>
<tr><td>**E-Id.**</td><td>$RR_{I,\{POL:negative,\ TNS:present\},\{UL\}}(\langle X, \sigma\rangle) = \langle Y\hat{\partial}l, \sigma\rangle$,
where Y is the realization of $\langle OK, \sigma\rangle$.</td><td></td></tr>
<tr><td>**E-Ie.**</td><td>$RR_{I,\{POL:negative,\ TNS:2nd\ past\},V}(\langle X, \sigma\rangle) = \langle [Z\ \underline{Y}], \sigma\rangle$,
where Y is the realization of $\langle UL, \sigma/\{TNS:present\}\rangle$, $Nar_I(\langle X,\sigma'\rangle) = \langle Z,\sigma'\rangle$, and $\sigma' = \sigma/\{POL:affirmative\}$.</td><td>(instead of **W-Ie**)</td></tr>
<tr><td colspan="2">**Block III**</td><td></td></tr>
<tr><td>**E-IIIa.**</td><td>$RR_{III,\{TNS:2nd\ past,\ AGR:\{PER:1,\ NUM:sg\}\},V}(\langle X, \sigma\rangle) = \langle Xam, \sigma\rangle$.</td><td>(= **W-IIIc**)</td></tr>
<tr><td>**E-IIIb.**</td><td>$RR_{III,\{TNS:2nd\ past,\ AGR:\{PER:2,\ NUM:sg\}\},V}(\langle X, \sigma\rangle) = \langle Xat, \sigma\rangle$.</td><td>(= **W-IIId**)</td></tr>
<tr><td>**E-IIIc.**</td><td>$RR_{III,\{TNS:2nd\ past,\ AGR:\{PER:3,\ NUM:pl\}\},V}(\langle X, \sigma\rangle) = \langle X\hat{\partial}t, \sigma\rangle$.</td><td>(= **W-IIIe**)</td></tr>
<tr><td>**E-IIId.**</td><td>$RR_{III,\{AGR:\{PER:1,\ NUM:sg\}\},V}(\langle X, \sigma\rangle) = \langle Xm, \sigma\rangle$.</td><td>(= **W-IIIf**)</td></tr>
<tr><td>**E-IIIe.**</td><td>$RR_{III,\{AGR:\{PER:2,\ NUM:sg\}\},V}(\langle X, \sigma\rangle) = \langle Xt, \sigma\rangle$.</td><td>(= **W-IIIg**)</td></tr>
<tr><td>**E-IIIf.**</td><td>$RR_{III,\{AGR:\{PER:1,\ NUM:pl\}\},V}(\langle X, \sigma\rangle) = \langle Xna, \sigma\rangle$.</td><td>(= **W-IIIh**)</td></tr>
<tr><td>**E-IIIg.**</td><td>$RR_{III,\{AGR:\{PER:2,\ NUM:pl\}\},V}(\langle X, \sigma\rangle) = \langle Xda, \sigma\rangle$.</td><td>(= **W-IIIi**)</td></tr>
<tr><td>**E-IIIh.**</td><td>$RR_{III,\{AGR:\{PER:3,\ NUM:pl\}\},V}(\langle X, \sigma\rangle) = \langle Xt, \sigma\rangle$.</td><td></td></tr>
</table>

Many of these rules have identical counterparts in Western Mari; note, for example, that rules **E-IIIa** through **E-IIIg** match rules **W-IIIc** through **W-IIIi**. There are, however, two main points of contrast: first, by rules **E-Ie** and **E-Ia**, a verb's negative second-past realizations are built upon its affirmative gerundial stem rather than on its negative gerundial stem (contrast rule **W-Ie** in (15)); and second, a verb's negative second-past realizations are, according to rule **E-Ie**, periphrases consisting of a primary second-past stem and a present-tense realization of the copula UL. Each negative present-tense realization of UL is, by **E-Id**, the result of compounding the corresponding realization of the negative auxiliary OK with UL's stem *∂l*; cf. Tables 11 and 12. The realization rules relevant for the present-tense inflection of the negative auxiliary OK in (23) are **E-Ic**, and **E-IIId** through **E-IIIh**.

By the rules in (23), the periphrase [*kolen om∂l*] `I didn't die' is the realization of the cell $\langle KOL, \{POL:neg, TNS:2^{nd}\ past, AGR:\{PER:1, NUM:sg\}\}\rangle$ in KOL's syntactic paradigm. This association with the property set {POL:neg, TNS:2nd past, AGR:{PER:1,

---

[31] We also assume that vowel hiatus is avoided by the elision of the second of two adjacent vowels (as in *ona-l* `we are not', i.e. *o-na-∂l*), and that intervocalic obstruents are subject to lenition (as in *o•a* and *o•∂t*, the second- and third-person plural present-tense realizations of the negative auxiliary OK).

[32] Rule **E-Ic** optionally applies more generally, in all plural affirmative present forms.

NUM:sg}} is transmitted from the periphrase [*kolen omǝ̂l*] to the verb phrase [$_{VP}$ *kolen omǝ̂l*] by the rule (16) at the morphology-syntax interface. It is this rule, in concert with the realizational morphology of Eastern Mari, that determines the content of this verb phrase; the usual principles of property unification to which c-structures are ordinarily subject are here overridden.

Comparable phenomena are widely observable. Consider a second example. In Udmurt, the imperfective past tense is a compound tense used to describe "a protracted or repeated activity occurring in the ... distant past" (Csúcs 1990:51). This tense is realized by the periphrastic combination of a future-tense form (inflected for subject agreement) with the invariant past form *val* of the copula, as in Table 13; compare the future-tense forms in Table 14.

TABLE 13. Affirmative imperfective past-tense realizations of Udmurt MÏNÏ `go'
[data from Suihkonen 1995:302[33]]

SG.	1	*mïno val* `I used to go (long ago)'
	2	*mïnod val*
	3	*mïnoz val*
PL.	1	*mïnom(ï) val*
	2	*mïnodï val*
	3	*mïnozï val*

TABLE 14. Affirmative future-tense realizations of Udmurt MÏNÏ `go'
[data from Csúcs (1988:142)]

SG.	1	*mïno* `I will go'
	2	*mïnod*
	3	*mïnoz*
PL.	1	*mïnom(ï)*
	2	*mïnodï*
	3	*mïnozï*

Neither part of an imperfective past-tense periphrase such as [*mïno val*] carries any exponent of an aspectual property such as durativity or habituality; yet, such a property is associated with the verb phrase [$_{VP}$ *mïno val*] as a whole. Moreover, while the finite head of [*mïno val*] is marked for future tense, the periphrase as a whole expresses the distant past tense. This departure from pure compositionality is, we claim, determined by the morphology of Udmurt: the temporal and aspectual properties of the verb phrase [$_{VP}$ *mïno val*] aren't deducible from the properties of its individual syntactic atoms by means of ordinary unification, but are instead the effect of a morphological rule of periphrasis realizing certain cells in the syntactic paradigm of MÏNÏ `go'. (This rule is formulated in (25) below.)

---

[33] Though the source for paradigm is Suihkonen 1995, for consistency we utilize the orthography used in various works by Csucs.

4.4 Distributed exponence in periphrases[34]

Though instances of extended exponence are far from rare, there is a tendency for each of the morphosyntactic properties realized by an inflected word form to have no more than a single exponent in that form's morphology. This tendency toward DISTRIBUTED EXPONENCE receives its fullest expression in heavily agglutinating languages. For instance, the Swahili verb form *ha-tu-ta-taka* `we will not want' has exactly one affixal exponent for each of the morphosyntactic properties it expresses: *ha-* expresses the property POL:neg; *tu-*, the property AGR:{PER:1, NUM:pl}; and *ta-*, the property TNS:future.

Periphrases likewise often exhibit distributed exponence. In Udmurt, for instance, the negative future-tense realizations of MÏNÏ `go' are periphrastic combinations of a realization of the negative verb U with a "connegative" form of MÏNÏ, as in Table 15; the connegative form realizes number but not person, while the negative verb form realizes person but not number (except in the first person, where the negative verb forms *ug* and *um* apparently express both person and number). Because a verb's imperfective past-tense realizations are built upon its future-tense realizations (cf. again Tables 13 and 14), the negative imperfective past-tense realizations of MÏNÏ `go' in Table 16 embody this same distribution of exponence.

TABLE 15. Negative future-tense forms of Udmurt MÏNÏ `go'
[data from Csúcs (1988:143)]

SG.	1	*ug mïnï*
	2	*ud mïnï*
	3	*uz mïnï*
PL.	1	*um mïne(le)*
	2	*ud mïne(le)*
	3	*uz mïne(le)*

TABLE 16. Negative imperfective past-tense realizations of Udmurt MÏNÏ `go'
[data from Suihkonen 1995:302]

SG.	1	*ug mïnï val*
	2	*ud mïnï val*
	3	*uz mïnï val*
		`s/he didn't used to go (long time ago)'
PL.	1	*um mïne(le) val*
	2	*ud mïne(le) val*
	3	*uz mïne(le) val*

We therefore propose distributed exponence as a third sufficient criterion for the identification of periphrases:

---

[34] It seems to us reasonable to posit as an additional sufficient criterion the phenomenon of multiple exponence whereby the same morphosyntactic property set(s) receives expression several times within the grammatical word. See Anderson 2001 for evidence concerning the existence of multiple exponence within synthetic wordforms and Sells (this volume) for its extension to periphrastic expressions.

(24) Criterion III: If the morphosyntactic property set associated with an analytic combination C has its exponents distributed among C's parts, then C is a periphrase.

In a purely syntactic approach to periphrasis, there is no particular reason to expect that periphrases should exhibit a comparable tendency toward distributed exponence; word combinations in syntax are, after all, sometimes highly redundant in their expression of shared morphosyntactic properties (as in the Swahili sentence *ki-kapu ki-kubwa ki-moja ki-lianguka* `one large basket fell', every one of whose words carries an exponent of the subject's gender and number). But if the economy of inflectional exponence exhibited by heavily agglutinating languages is seen as a property of rules which define morphological paradigms, then the assumption that periphrases are morphological in their definition entails that periphrases should be no less likely to exhibit this same economy.

The analysis of Urdmurt verb morphology in (25) accounts both for the noncompositionality exemplified in Table 13 and for the distribution of exponence exemplified in Tables 15 and 16.

(25) Some realization rules for future-tense and imperfective past-tense realizations in Udmurt

Block I

U-Ia. $RR_{I,\{\text{TNS:distant past, ASP:durative}\},V[\text{nonauxiliary}]}(\langle X,\sigma\rangle) = \langle[\underline{Y}\ val],\sigma\rangle$, where Y is the realization of $\langle X,\sigma/\{\text{TNS:future, ASP:simple}\}\rangle$.

U-Ib. $RR_{I,\{\text{POL:aff, TNS:fut}\},V[\text{nonauxiliary}]}(\langle X,\sigma\rangle) = \langle Xo,\sigma\rangle$.

U-Ic. $RR_{I,\{\text{POL:neg, AGR:\{NUM:sg\}\}},V[\text{nonauxiliary}]}(\langle X,\sigma\rangle) = \langle X\ddot{\imath},\sigma\rangle$.

U-Id. $RR_{I,\{\text{POL:neg, AGR:\{NUM:pl\}\}},V[\text{nonauxiliary}]}(\langle X,\sigma\rangle) = \langle Xe(le),\sigma\rangle$.

Block II

U-IIa.[35] $RR_{II,\{\text{POL:neg}\},V[\text{nonauxiliary}]}(\langle X,\sigma\rangle) = \langle[\underline{Y}\ X],\sigma\rangle$, where Y is the realization of $\langle U,\sigma\rangle$.

Block III

U-IIIa. $RR_{III,\{\},\{U\}}(\langle X,\sigma\rangle) = \langle Xg,\sigma\rangle$.

U-IIIb. $RR_{III,\{\text{POL:aff, AGR:\{PER:1, NUM:pl\}\}},V}(\langle X,\sigma\rangle) = \langle Xm,\sigma\rangle$.

U-IIIc. $RR_{III,\{\text{POL:aff, AGR:\{PER:2\}\}},V}(\langle X,\sigma\rangle) = \langle Xd,\sigma\rangle$.

U-IIId. $RR_{III,\{\text{POL:aff, AGR:\{PER:3\}\}},V}(\langle X,\sigma\rangle) = \langle Xz,\sigma\rangle$.

Block IV

U-IVa. $RR_{IV,\{\text{POL:affirmative, AGR:\{NUM:pl\}\}},V[\text{nonauxiliary}]}(\langle X,\sigma\rangle) = \langle X\ddot{\imath},\sigma\rangle$.

Rule **U-Ia** causes the cell <MÏNÏ, {POL:aff, TNS:distant past, ASP:durative, AGR:{PER:2, NUM:pl}}> in the syntactic paradigm of MÏNÏ `go' to be realized as the periphrase [*mïnodï val*], neither of whose parts is itself an expression of durative aspect and whose head is,

---

[35] We assume here that the Udmurt negative verb U is subject to a rule of paradigm linkage analogous to the rule for Western Mari AK in (19).

on its own, an expression of the future tense; at the morphology/syntax interface, principle (16) associates this periphrase with a c-structure in which the usual, compositional pattern of property unification is suspended. Moreover, rule **U-IIa** causes <MÏNÏ, {POL:negative, TNS:distant past, ASP:durative, AGR:{PER:2, NUM:pl}}> to be realized as the periphrase [[*ud mïne(le)*] *val*], whose exponence is distributed: the head *ud* is (by **U-IIIc**) an exponent of person but not number, while its nonhead element *mïne(le)* is an exponent of number but not person.

4.5 Some confirming evidence

Here we examine two types of evidence confirming the need for a morphological approach to periphrasis: the close connection between periphrastic realizations and synthetic realizations in language change, and the participation of periphrasis in generalizations about morphological markedness.

4.5.1 Periphrasis and language change

A widely observed phenomenon in historical linguistics is the diachronic development of periphrasis into synthetic morphology. This phenomenon follows very naturally from the conception of periphrasis advocated here. Periphrasis is, in this approach, just one more type of morphological exponence. The development of synthetic morphology from periphrasis is therefore not different, in principle, from the development of fusional morphology from agglutination: both sorts of developments involve an increasing degree of fusion in the inflectional realization of a paradigm's cells. Our approach predicts that just as one can observe different degrees of progress in the development from agglutination to fusionality, one should likewise find different degrees of progress in the development of synthesis from periphrasis; Mari itself provides compelling evidence of this sort of gradation.

Consider again the synthetic negative second-past realizations in Western Mari (given above in Table 7). It is clear that these descend historically from periphrases--in particular, from periphrastic combinations of the negative gerund with affirmative present-tense forms of the copula UL; compare the free forms of this copula in Table 17.

TABLE 17. Affirmative present-tense realizations of the Mari *am*-conjugation copula UL (Western dialect) [Alhoniemi 1985:114]

SG.	1	*ə̂l-a-m*
	2	*ə̂l-a-t*
	3	*ə̂l-eš*
PL.	1	*ə̂l-ə̂-na*
	2	*ə̂l-ə̂-da*
	3	*ə̂l-ə̂-t*

This development from periphrasis to synthesis in the second past is even more extensively observable in Northwest Mari. Consider the affirmative second-past realizations of TOL `go' in Table 18.

TABLE 18. Affirmative second-past realizations of the Mari *am*-conjugation verb TOL `go' (Northwest dialect) [Bereczki (1990:55)]

SG.	1	*tol-ə̂n-l-am* `I went'
	2	*tol-ə̂n-l-at*
	3	*tol-ə̂n*
PL.	1	*tol-ə̂n ulna*
	2	*tol-ə̂n ulda*
	3	*tol-ə̂n ult*

Here, the plural realizations consist of an uninflected affirmative gerund with an independent copula inflected for present tense and subject agreement. By contrast, this copular construction has become synthetic in the first- and second-person singular realizations in Table 18; like the negative second-past realizations in Western Mari, the singular affirmative realizations in Table 18 now involve a gerundial stem whose absolute form appears in the third-person singular and whose conjunct form (suffixed with *l*) otherwise appears with the appropriate person/number marker. The development from periphrasis to synthesis has progressed partway across the affirmative second-past paradigm.

In a morphological approach to periphrasis, the affirmative second-past realizations of Northwest Mari in Table 18 and their Western Mari counterparts in Table 7 express the same syntactic paradigm; that is, the relevant different between these dialects is a difference not in the inventory of cells available for realization but in the morphological rules by which these cells are realized. In particular, the Northwest dialect's system of realization rules differs from the system in (15) in two relevant ways. First, it has an additional rule of periphrasis whose effect is to license periphrases in plural realizations of the affirmative second past; second, it employs a single rule to define the default primary stem of the second past. These rules are formulated in (26):

(26) Realization rules for second-past verb realizations (Northwest dialect)

Block I (as in (11) except that **NW-Ia** takes the place of **W-Ic** and **W-Ie**)
NW-Ia $RR_{I,\{POL:\alpha,\ TNS:2nd\ past\},V}(\langle X, \sigma \rangle) = \langle Y, \sigma \rangle$, where Y is an absolute stem form which

(a) is identical in form to the realization of $\langle X, \{VFORM:gerund; POL:\alpha\} \rangle$ and
(b) has Y*l* as its conjunct stem form.

Block II (as in (11) except for the additional rule **NW-IIa**)
NW-IIa $RR_{II,\{POL:affirmative,\ TNS:2nd\ past,\ AGR:\{NUM:pl\}\},V}(\langle X, \sigma \rangle) = \langle [X\ \underline{Y}], \sigma \rangle$, where Y is the realization of $\langle UL, \sigma/\{TNS:present\} \rangle$.

Block III (as in (11))

It is a virtue of the morphological approach to periphrasis that it allows the Western/Northwest contrast in the affirmative second past to be so simply localized in the morphology.

The Northwest forms in Table 18 suggest a change in progress: gradually, the copulative verb used in the periphrastic expression of the affirmative second past has become enclitic, then reanalyzed as synthetic morphology. A similar change in progress is actually documented in Eastern Mari, where the plural realizations of the copula UL in the negative present tense appear sometimes as single-word forms (as in Table 11), but sometimes as periphrases (as in Table 19); presumably the periphrastic realizations are losing ground among innovative speakers.

TABLE 19. Variation in negative present-tense realizations of the Mari *am*-conjugation copula UL (Eastern dialects) [Alhoniemi 1985:111,116]

SG.	1	*om-ə̂l* `I am not'
	2	*ot-ə̂l*
	3	*oɣ-ə̂l*
PL.	1	*ona-l ~ oɣə̂na-l ~ oɣə̂na ul*
	2	*oδa-l ~ oɣə̂δa-l ~ oɣə̂δa ul*
	3	*oɣə̂t-ə̂l ~ oɣə̂t ul*

4.5.2 Periphrasis and morphological markedness

Very often, the types of exponence observable among the realizations of a lexeme's paradigm correlate with the degree of markedness of the morphosyntactic property sets which those forms realize. For instance, exponents of more highly marked morphosyntactic properties tend to be less fusional (cf. Greenberg 1966, Mayerthaler 1988): in Sanskrit, for instance, the nominative singular is often expressed purely by stem gradation (e.g. PITAR `father', nom. sg. *pitā*, RĀJAN `king', nom. sg. *rājā*), while the dative singular is always expressed suffixally (*pitr-e* `to the father', *rājñ-e* `among the kings'); in Swahili, negation and first-person singular subject agreement are expressed by a portmanteau prefix *si-* (e.g. *si-ta-taka* `I will not want') while negation and first-person plural subject agreement are expressed separately by the respective suffixes *ha-* and *tu-* (*ha-tu-ta-taka* `we will not want'); and so on.

In a theory in which periphrasis is regarded as a kind of morphological exponence, periphrastic exponence would be expected to participate in this same correlation. In particular, since periphrasis is by definition nonfusional, one would expect that in paradigms in which synthesis and periphrasis exist side by side, the incidence of periphrasis will be associated with more highly marked morphosyntactic properties. And such is indeed overwhelmingly the case. For instance, among the Northwest dialect realizations in Table 18, it is those whose property sets include the marked number specification `plural' which preserve periphrasis; similarly, periphrasis is the default expression of the marked polarity specification `negative' in Western Mari. This phenomenon is dramatically evident in the Samoyedic language Tundra Nenets as illustrated in Table 20.

TABLE 20. Declension of *ti* `reindeer' in Tundra Nenets
(Data from Salminen 1997)

		Singular	Dual	Plural
Grammatical cases	Nominative	*ti*	*tex^{o}h*	*tiq*
	Accusative	*tim*	*tex^{o}h*	*tí*
	Genitive	*tih*	*tex^{o}h*	*tíq*
Local cases	Dative	*ten^{o}h*	*tex^{o}h nyah*	*tex^{o}q*
	Locative	*tex^{o}na*	*tex^{o}h nyana*	*tex^{o}qna*
	Ablative	*texød^o*	*tex^{o}h nyad^o*	*texøt^o*
	Prosecutive	*tew^{o}na*	*tex^{o}h nyamna*	*teqm^{o}na*

Periphrastic expression in Nenets nominals, consisiting of the dual stem plus an appropriately case inflected form of the postposition nya, occurs solely for those cells which contain the most marked value for case (namely the local cases) as well as the most marked value for number (namely dual); all other morphosyntactic property sets receive synthetic expression.

A theory whose notion of exponence encompasses synthetic but not periphrastic markings affords no coherent articulation of the overarching generalization which such cases embody; in the grammatical ontology of such a theory, the phenomena which ought to be subsumed under this generalization--phenomena such as fusion, agglutination, periphrasis--fail to constitute any kind of natural class.

5. **Conclusions**

Drawing on the assumptions of inferential-realizational morphology, we have argued that certain periphrastic expressions are directly projected from morphological paradigms by realization rules. We have adduced criteria for distinguishing morphologically defined periphrases from ordinary syntactic complementation in Uralic. These criteria motivate the adoption of the Periphrastic Realization Hypothesis (given above in (9)). They also require a modification of the basic principle regulating the relation between morphology and syntax, permitting the exponence of lexical representations to be realized as independent and possibly discontinuous elements in c-structure. Finally, this type of analysis is facilitated by appealing to standard inferential-realizational assumptions concerning the strict separation of content from form and, in effect, represents a trivial extension of these assumptions so that they apply to periphrastic expressions. This is implemented here by interpreting lexical representations in terms of cells in syntactic and morphological paradigms which are put in correspondence by rules of paradigm linkage and by providing realization rules which account for the surface syntactic exponence of roots and stems in morphological paradigms.

Appendix: Remarks on the morphophonology of Mari

The morphophonology of Mari is quite complex; here we merely describe those rules relevant to the definition of the forms in (2)-(4). Among the vowels introduced by the realization rules in (11), some are subject to vowel harmony (those in **W-IIIa** through **W-IIIe**, **W-IIIh**, and **W-IIIi**) and some are not (those in **W-Ib**, **W-If**, **W-Ig**, **W-IIb**, **W-IIIl**, and **W-IIIm**). Moreover, a suffixal vowel which is subject to vowel harmony only harmonizes when the suffix joins with a stem whose last vowel is a trigger; vowels are triggers by default, but the vowel introduced by rule **W-Ig** is not a trigger (so that even though *kolenäm* `I went' exhibits vowel harmony, *kolə̂δelam* `I didn't go' does not).

The stops introduced by rules **W-Ib** and **W-IIIi** are subject to intervocalic lenition (as in *kolə̂neδä* `you (pl.) want to die' and *koleβə* `they died') and to obstruent voicing assimilation (as in *kolə̂šta* `you (pl.) died').

All of the rules in Block III which apply in the definition of negative second-past realizations select the conjunct primary stem (e.g. *kolə̂δel-*) over its absolute counterpart (*kolə̂δe-*).

In the definition of the negative verb AK's affirmative first-past realizations, the ə-ə hiatus occasioned by the application of rules **W-IIIa** and **W-IIIb** is eliminated by vowel coalescence.

We assume that all of the morphophonological rules described here have the status of morphological metageneralizations; see Stump (2001:47ff) for discussion.

References

Abondolo, Daniel, ed. 1998. *The Uralic Languages*. London & New York: Routledge.

Ackerman, Farrell 1984. Verbal modifiers as argument aking predicates: Complex verbs as predicate complexes in Hungarian. *Groningen Working Papers in Linguistics.* Groningen: University of Groningen. 23-71.

Ackerman, Farrell 1987. Miscreant Morphemes: Phrasal Predicates in Ugric. UC Berkeley Ph.D. disseration.

Ackerman, Farrell, and Gert Webelhuth. 1998. *A Theory of Predicates.* CSLI Publications.

Alhoniemi, Alho 1985. *Marin Kielioppi*. Helsinki: Suomalais-Ugrilainen Seura.

Alhoniemi, Alho 1988. Das Tscheremissische. In Sinor, ed., 1988, pp.84-95.

Alhoniemi, Alho 1993. *Grammatik des Tscheremissischen (Mari).* Hamburg: Helmut Buske Verlag.

Anderson, Stephen R. 1992, *A-Morphous Morphology*, Cambridge University Press.

Anderson, Stephen R. 2001. On some issues in morphological exponence. *Yearbook of Morphology 2000.* 1-18.

Andrews, A. 1990. Unification and morphological blocking. *Natural Language and Linguistic Theory.* 8(4), 507-557.

Aronoff, M. 1994. *Morphology by Itself: Stems and Inflectional Classes.* Cambridge: MIT Press.

Beard, Robert 1995, *Lexeme-Morpheme Base Morphology*, Albany: SUNY Press.

Bereczki Gábor 1990. *Chrestomathia Ceremissica*. Budapest: Tankönyvkiadó.

Blevins, J. 2001. Lexemic stems in Western Germanic. Ms. Cambridge University.

Blevins, J. 2000. 'Markedness and blocking in German declensional paradigms'. In B. Stiebels and D. Wunderlich (eds.) Lexicon in Focus, Studia Grammatica 45:83-103, Berlin: Akademie-Verlag.

Börjars, Kersti, Nigel Vincent and Carol Chapman 1997, Paradigms, periphrases and pronominal inflection: a feature-based account, in G. Booij and J. van Marle, (eds.), *Yearbook of Morphology 1996*, 155-180, Dordrecht: Kluwer.

Brassil, Daniel. 2001. Periphrasis: Standard Lexicalism and Realization Lexicalism. Ms.. UCSD.

Bresnan, Joan and Sam Mchombo 1995, The lexical integrity principle: evidence from Bantu, *Natural Language and Linguistic Theory* 13: 181-254.

Butt, M. et. al. 1996. Multilingual processing of auxiliaries in LFG. In D. Gibbon. ed. *Natural Language Processing and Speech Technology: Results of the 3rd KONVENS Conference.* Mouton de Gruyter. 111-122.

Carstairs-McCarthy, Andrew. 2000. Lexeme, word-form, paradigm. In Geert Booj, Christian Lehmann, Joachim Mugdan in collaboration with Wolfgang Kesselheim, Stavros Skopeteas Morphologie : ein internationales Handbuch zur Flexion und Wortbildung / Morphology : An international handbook on inflection and word-formation. Berlin ; New York : Walter de Gruyter. 595-607.

Csúcs, Sándor. 1987. *Chrestomathia Votiaca.* Budapest: Tankönyvkiadó

Csúcs, Sándor 1988. Die wotjakische Sprache. In Sinor, ed., 1988, pp.131-146.

Csúcs, Sándor 1998. Udmurt. In Abondolo, ed., 1998, pp.276-304.

Falk, Y. 1984. The English auxiliary system: A lexical-functional analysis. *Language* 60(3), 483-509.

Greenberg, Joseph H. 1966. *Language universals, with special reference to feature hierarchies*. The Hague: Mouton.

Gukhman, M. M. 1955. Glagolnye analaticheskie konstruksii kak osobyj sochetanij chastichnogo I polnogo slova (na materiale istorii nemetskogo iazyka). In V. V. Vonogradov ed. Voprosy grammatischeskogo stoia. Moscow: Academic Sciences.

Gukhman, M. M. 1963. Kriterii vydelenia glagolnikh analaticheskikh konstruksij iz drugikh tipov slovosochetanij. In V. M. Zhirmunskij and O. P. Sunik eds. Morfologicheskaia struktura slova v iazykach razlichnykh tipov. Moscow: Academic Sciences. 199-202.

Haspelmath, Martin. 2000. Periphrasis. In In Geert Booj, Christian Lehmann, Joachim Mugdan in collaboration with Wolfgang Kesselheim, Stavros Skopeteas Morphologie : ein internationales Handbuch zur Flexion und Wortbildung / Morphology : An international handbook on inflection and word-formation. Berlin ; New York : Walter de Gruyter. 654-664.

Kangasmaa-Minn, Eeva 1998. Mari. In Abondolo, ed., 1998, pp.219-248.

Kel'makov, Valentin and Sara Hännikäinen. *Udmurtin Kielioppia ja Harjoituksia.* Helsinki: Suomalais-Ugrilainen Seura.

Koenig, Jean-Pierre. 1999. Lexical Relations. Stanford: CSLI publications.

Matthews, P. H. 1972, *Inflectional Morphology: A Theoretical Study Based on Aspects of Latin Verb Conjugation*, Cambridge University Press.

Matthews, P. H. 1991, *Morphology* [2nd edn.], Cambridge University Press.

Mayerthaler, Willi. 1988. *Morphological naturalness*. Ann Arbor: Karoma.

Mitchell, E. 1993. Morphological evidence for syntactic structure: the Finno-Ugric languages and nglish. PhD. dissertation, Cornell University.

Mohanan, K. P. 1982. Grammatical relations and clause structure in Malayalam. In J. Bresnan ed. *The Mental Representation of Grammatical Relations*. Cambridge: MIT Press. 504-589.

Niño, Maria-Eugenia. 1997. The multiple expression of inflectional information and grammatical architecture. In Francis Corblin, et,. al. eds. *Empirical Issues in Formal Syntax and Semantics*. 127-47. Berne: Peter Lang.

Orgun, O. 1996. Sign-based morphology and phonology with special attention to Opimality Theory. PhD. dissertation, UC Berkeley.

Riehemann, S. 2000. Type-Based Derivational Morphology. Journal of Comparative Germanic Linguistics 2: 49-77.

Robins, R. H. 1959. In defence of WP. *Philological Society of London.* 116-144.

Sadler, L & A. Spencer. 2001. Syntax as an exponent of morphological features *Yearbook of Morphology 2000.* 71-97.

Salminen, Tapani. 1997. *Tundra Nenets inflection.* Suomalias-Ugrilainen Seura. Helsinki. Sebeok, Thomas, and Frances J. Ingemann. 1961. *An eastern Cheremis manual: phonology, grammar, texts and glossary.* Bloomington: Indiana University publications [Uralic and Altaic series, v. 5].

Sells, Peter. Syntactic information and its morphological expression. Ms. Stanford University.

Serebrennikov, V. A. 1964. *Osnovye linii razvitija padežnoj I glagolnoj sistem v uralskix jazykax.* Nauk, Moscow.

Sinor, Denis, ed. 1988. *The Uralic Languages: Description, History and Foreign Influences*. Leiden: E. J. Brill.

Spencer, A. 2001. The Word-and-Paradigm approach to morphosyntax'. Transactions of the Philological Society, 99:279-313.

Spencer, A. to appear. Peiphrastic Paradigms in Bulgarian. Ms. University of Essex.

Stump, Gregory T. 1993. On rules of referral, *Language* 69: 449-479.

Stump, Gregory T. 2001. *Inflectional Morphology*. Cambridge University Press.

Stump, Gregory T. 2002. Morphological and syntactic paradigms: Arguments for a theory of paradigm linkage. *Yearbook of Morphology*.

Suihkonen 1995. Udmurt-English-Finnish Dictionary with A Basic Grammar of Udmurt. Suomalais-Ugrilainen Seura. Helsinki.

Tsypanov, J. 2001. Analyyttisten aikamuotojen systeemi komin kielessa. In T. Seilenthal et.al. eds. Congressusnonus internationalis Fenno-Ugristarum. Tartu 321-324.

Zhirmunskij, V. M. 1963. O granitsakh slova. In V. M. Zhirmunskij and O. P. Sunik eds. Morfologicheskaia struktura slova v iazykakh razlichnykh tipov. Moscow: Nauk. 3-33.

Zorina, Z. G. et.al. 1990. Mariiskii Iazyk dl'a vsekh. Mariiskoe Knizhnoe Izdatelstvo. Ioshkar-Ola.

Zwicky, Arnold M. 1985, How to describe inflection, in Mary Niepokuj, Mary Van Clay, Vassiliki Nikiforidou and Deborah Feder (eds.), *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*, 372-386, Berkeley Linguistics Society.

Zwicky, Arnold M. 1990, Inflectional morphology as a (sub)component of grammar, in Wolfgang U. Dressler, Hans C. Luschützky, Oskar E. Pfeiffer and John R. Rennison (eds.), *Contemporary Morphology*, 217-236, Berlin: Walter de Gruyter.