Morphological phrasemes and Totonacan verbal morphology*

DAVID BECK AND IGOR MEĽČUK

Abstract

The existence of restricted or phraseologized complex expressions such as clichés, collocations, and idioms (collectively known as phrasemes) is well-known and widely accepted in the domain of multi-word expressions. What is not so widely recognized is the existence of the same type of phraseologized expression at the morphological level — restricted complex morphological expressions, or morphological phrasemes. Morphological phrasemes, found in both derivation and inflection, are governed by the same principles of phraseologization that govern phrasemes at the lexical-syntactic level, and have roughly the same subtypes, including morphological collocations and morphological idioms. In addition to offering clear advantages in terms of useful descriptive practice and formal economy, the recognition of phraseologized morphological expressions makes clear the parallels between linguistic signs at the lexical and morphemic level, this isomorphism falling out from the conventionalized nature of the mapping between linguistic meaning and linguistic form. Contrary to many current approaches to morphological theory that reject the morpheme as a meaning-bearing element in the structure of words, this paper argues that the existence of conventionalized uses of sublexical elements gives strong support to the utility of the morpheme as unit of linguistic analysis.

1. Introduction: The problem stated

A perennial set of problems in the description of morphological expressions (that is, any syntagmatic combination of linguistic signs that does not exceed the boundaries of a word) is centered around cases in which the formal description of a particular morphological string departs in some way from the canonical one-to-one pairing of its individual analyzable parts — roughly speaking, morphs — to individual meanings. This paper is concerned with a particular

type of departure from the canonical pattern, illustrated by the present indicative form of the Upper Necaxa Totonac (UNT) verb shown in (1c):[1]

(1) a. *ikłtatá:*
         ik–łtatá:–ya:[2]
         1SG$_{SUB}$–sleep–INCOMP
         'I sleep' [1SG form; = ŁTATÁ:$_{INCOMP, 1SGSUB}$]

     b. *łtata:yá:w*
         łtatá:–ya:–w
         sleep–INCOMP–1PL$_{SUB}$:INCL
         'we-including-you sleep' [1PL inclusive form;
         = ŁTATÁ:$_{INCOMP, 1PLSUB:INCL}$]

     c. *ikłtata:yá:w*
         ik–łtatá:–ya:–w
         1SG$_{SUB}$–sleep–INCOMP–1PL$_{SUB}$:INCL
         'we-excluding-you sleep' [1PL exclusive form;
         = ŁTATÁ:$_{INCOMP, 1PLSUB:EXCL}$]

In (1a), the prefix *ik-* marks the first-person singular subject, and in (1b), the suffix *-w* marks the first-person plural inclusive subject. However, in (1c), neither of these affixes contributes to the word form precisely the same meaning as in the previous examples; instead, the two affixes taken together mark a different subject — the first-person plural exclusive. In this case, and in many similar instances, it seems impossible to systematically map meaning to form in a one-to-one manner, or to assign a unique elementary marker to the first-person plural exclusive subject. Facts such as these have been used as arguments against the recognition of affixes or, more generally, morphs and the morphemes they belong to as legitimate units of linguistic description — that is, as entities that must be included in some independent form in the lexicon as discrete elements that carry meanings (e.g., Anderson 1992; Stump 2001). In a variety of such theories of morphology, often grouped together under the heading "word-based approaches" (e.g., Matthews 1991; Zwicky 1992; Aronoff 1994; Beard 1995; Spencer 2001; Blevins 2003), it is considered unnecessary, even misguided, to try to assign meanings to affixes like *ik-* and *-w* in any of the examples in (1), the basic descriptive unit of morphology being not morphs or morphemes, but the inflected word form as a whole and/or the lexeme to which that word form belongs.

There is, however, another approach that might be taken to the data in (1). Rather than abandoning the concept of the morpheme as a meaningful element and losing the intuitive analysis that *ik-* is the marker of the first-person singular subject and *-w* is a marker of the first-person plural inclusive subject, an alternative would be to say that these affixes have these meanings in the language but lose them in specific, well-defined circumstances of use, such as those in

(1c). In other words, these markers can be assigned specific individual meanings, but when they come together, the meaning of the combination is different from the regular sum of these individual meanings. On the level of phrases (that is, multi-word expressions) this kind of treatment is routine and constitutes a familiar linguistic phenomenon — that of phraseologized expressions, or *phrasemes* (Mel'čuk 1995). The existence of such phraseologized combinations of elements is a well-known and important property of natural languages, and phrasemes such as collocations and idioms have been intensely studied in the domain of lexical-syntactic expressions, or phrases. They are much less studied in morphology, although the parallel between phraseologized expressions on the lexical-syntactic and morphological levels has been noted on more than one occasion in the past (Pike 1961; Mel'čuk 1964, 1982: 118, 1993–2000: vol. 4, 1995, 2006a; Aronoff and Sridhar 1984; Beck 2007). These parallels will be the focus of this paper. Using for the most part (although not exclusively) data from Upper Necaxa Totonac, a language from a family traditionally considered to be highly agglutinating with strictly compositional morphology, we will argue that phraseologization is not restricted to a particular type of linguistic sign — that is, not just to phrases. The defining characteristics of phrasemes, paradigmatic restrictedness and syntagmatic noncompositionality, characterize in principle all types of complex linguistic sign. Therefore, word forms built out of morphemes show the same properties of phraseologization as phrases built out of lexemes. Just as the constituent units of phrasal phrasemes (that is, lexemes) must be recognized as meaningful units in their own right, so must the constituents of morphological phrasemes — morphemes.[3]

Thus, this paper pursues two complementary ends. On the one hand, we attempt to account for some interesting and problematic Totonacan data from the realm of verbal inflection in terms of morphological phrasemes. In order to accomplish this, we propose a systematization of different types of morphological phrasemes, including definitions, new terminology, and formal means of representation. This is, so to speak, our second, theoretical goal. In so doing, we hope to both offer an economical description of a number of phraseologized combinations of morphemes in Totonacan, and to develop and sharpen our descriptive tools in the process, drawing at times on data from other languages when the need arises. It might perhaps have been more elegant to write a purely theoretical paper, or to concentrate only on the subset of phenomena exemplified by Totonacan inflection — but it seems to us that our hybrid approach is in fact the best possible way for linguistic research (apart from paper writing) to proceed, through the development of theoretical proposals which take as a starting point a particular set of novel or unusual data.

In the remainder of this paper we will proceed as follows. Section 2 sketches out a general theory of phraseology and illustrates it with more or less familiar examples of phraseologized multi-word expressions; in Section 3 we discuss

various types of morphological expression in terms of this theory, bringing our terminology to bear on the Totonacan data that initially drew our attention to this problem. Finally, in Section 4 we consider some of the implications of our findings for morphological phraseology and for linguistics in general.

2.   Phraseologization and phrasemes

One of the central properties of natural language is the existence of *non-free* complex linguistic signs — that is, complex signs such that not all of their constituent elements are chosen freely by the Speaker based on their meanings and the meaning the Speaker wishes to express,[4] but which are instead bound by various types of contextual (linguistic, pragmatic, or conventional/usage-based) restrictions. The class of non-free signs subsumes the full range of *phraseologized expressions*, including pragmatemes, collocations, clichés, and idioms, all of which contain one or more constituent elements whose choice is non-free, or *restricted*.

Free complex signs are illustrated by the phrasal expressions in (2):

(2)   a.   to **consider** the **chances** of **winning**
       b.   to **estimate** the **probability** of **victory**
       c.   to **evaluate how likely** it is to **win**
       d.   to **ponder** the **likelihood** of **being victorious**

For a phrase, "being free" means two things:

(i)   *Paradigmatic freedom*: Each of the three full constituent signs in (2a) — that is, the boldfaced portion of the phrase, which excludes the structural, or grammatical, words — is chosen by the Speaker for its meaning without any individual restriction. Each can be replaced by any (more or less) synonymous sign and the resulting phrase, such as those in (2b)–(2d), will be no less acceptable than (2a). Phrases (2a)–(2d) are therefore *unrestricted*.

(ii)   *Syntagmatic freedom*: All of the constituent signs in (2) are combined according to the general rules of English grammar, so that phrases (2a)–(2d) are fully regular. In particular, these phrases are (*semantically*) *compositional*: their meaning is a regular sum of the meanings of their constituents.

Thus, a *free* complex sign is unrestricted and compositional. If either or both of these defining properties — that is, paradigmatic freedom (unrestrictedness) or syntagmatic freedom (compositionality) — is violated, the sign is non-free, or phraseologized; phraseologized signs are called *phrasemes*. Using these prop-

erties, it is possible to define the three major types of phrases found in natural languages. As shown in Table 1, there are potentially four types of phrases defined in this way. The first type, the unrestricted compositional phrase, corresponds to the free phrase. If the properties of unrestrictedness and compositionality were logically independent, we would be left with three additional types of phrase: restricted compositional, restricted noncompositional, and unrestricted noncompositional. However, unrestricted noncompositional phrases are logically impossible. If a phrase is noncompositional, it has to be specified as a whole — that is, it has to be learned and stored as a single "chunk" independently of the meanings of its constituents; all of its constituents taken together form a lexical unit and none of its constituents are freely chosen. Therefore, it is restricted. This means that phrasemes fall into only two major types: 1) compositional phrasemes and 2) noncompositional phrasemes, or idioms. The overall defining feature of phrasemes is thus their restrictedness: all phrasemes are restricted complex linguistic signs, whereas some are compositional, and others noncompositional.

Within the two types of phraseme defined in this way, we can further distinguish various subtypes, given in Table 2. As this table shows, the property of compositionality serves to distinguish collocations and other compositional phrases from idioms.[5] In the remainder of this section, we will define each of these subtypes of phraseme and expand upon certain key concepts needed for their definition: we begin with compositional phrasemes and a discussion of

Table 1.   *Major types of phrases*

syntagmatic freedom paradigmatic freedom	COMPOSITIONAL	NONCOMPOSITIONAL
UNRESTRICTED	free phrases	[impossible]
RESTRICTED (= PHRASEMES)	pragmatemes, collocations, and clichés	idioms

Table 2.   *Types and subtypes of phrasemes*

compositional	pragmateme collocation cliché
noncompositional	weak idiom semi-idiom strong idiom

the notion of compositionality (Subsection 2.1), and then move on to noncompositional phrasemes (2.2). In order to facilitate the discussion, we will confine ourselves to speaking in terms of binary phrasemes (phrasemes containing just two constituents), though whatever is said is easily generalizable to larger expressions. We will also limit ourselves in the next two subsections to discussing only *phrasal* phrasemes (that is, phrasemes built from words), as these are more familiar and this will facilitate the exposition. We will reserve discussion of morphological phrasemes (phrasemes built of morphemes), the thrust of this paper, until Section 3, once the necessary concepts have been well established.

2.1.   *Compositional phrasemes*

Compositionality is the most important syntagmatic characteristic of complex linguistic signs (Mel'čuk 2004). It is defined in terms of the operation of linguistic union $\oplus$ (Mel'čuk 2006a: 366). The nature of this operation is determined in language L by a language-specific set of rules that describe the combination of signs and their components in L. A complex sign AB is compositional if and only if $AB = A \oplus B$, which means that

1) 'AB' = 'A' $\oplus$ 'B' (the meaning of AB is a regular sum of the meanings of A and B);
2) /AB/ = /A/ $\oplus$ /B/ (the form of AB is a regular sum of the forms of A and B); and
3) $\Sigma_{AB} = \Sigma_A \oplus \Sigma_B$ (the syntactics, or combinatorial properties, of AB is a regular sum of the syntactics of A and B).

To simplify things, we allow ourselves to concentrate only on semantic compositionality — that is, 'AB' = 'A' $\oplus$ 'B'. The free phrase *kick the ball* is semantically compositional ('kick the ball' = 'kick' $\oplus$ 'ball' $\oplus$ 'the'), and the meaning of the whole expression is therefore predictable from the meanings of its parts. On the other hand, the notorious set phrase ˹KICK THE BUCKET˺ is noncompositional ('kick the bucket' ≈ 'die' ≠ 'kick' $\oplus$ 'bucket' $\oplus$ 'the').[6]

The term *compositional* as it used here is distinct from another common use of the term as a synonym for what we would refer to as *transparent* (see, for example, Langacker [1987, 1991]; Jackendoff [2002]; Schultze-Berndt [2000]; Seifart [2005]). The transparency of a complex sign for an Addressee is the degree of psychological closeness (by analogy, metaphor, or metonymy) either between the usual meaning of a sign and its contextually-restricted meaning, or between the sum of the usual meanings of the constituents of an idiom and the idiom's conventional meaning. Clearly, transparency in this sense is an important concept in the study of the origins of phraseologized expressions and the psychological relations between such expressions and their components. How-

ever, transparency is not the same as compositionality: a fairly transparent expression can be noncompositional (e.g., ⌐STUFF ONE'S FACE⌐ 'eat prodigiously, to the point of satiation'). Transparency is also a matter of degree, whereas compositionality, as we use the term here, is a binary distinction: AB either is or is not the linguistic union of A and B. An idiom such as ⌐SITTING DUCK⌐ or ⌐MOVE HEAVEN AND EARTH⌐ is more transparent and thus much easier to understand (even for non-native speakers) than ⌐SHOOT THE BREEZE⌐ or ⌐HAIR OF THE DOG THAT BIT X⌐; yet from the point of view of semantic compositionality, all four are equivalent: they are noncompositional (i.e., idioms). The transparency of a complex sign is relevant for, among other things, the interpretation of texts by actual speakers, and depends on psychological and psycholinguistic factors. Compositionality is more relevant to the formal linguistic or computational modeling of text synthesis.

From this perspective, compositionality is a formal feature of the linguistic representation of signs — more specifically, it is a characteristic of a certain type of mapping between form and meaning. Thus, for a particular phrasal complex sign AB, whose meaning 'S' maps onto the configuration of lexemes $A \oplus B$ (that is, 'S' $\Leftrightarrow A \oplus B$), we say that it is compositional if and only if 'S' can be decomposed into meanings 'S$_1$' and 'S$_2$' ['S' = 'S$_1$' $\oplus$ 'S$_2$'] such that 'S$_1$' $\Leftrightarrow A$ and 'S$_2$' $\Leftrightarrow B$. Such meaning-to-form mappings can be, and often are, psychologically transparent for speakers of the language, but just as often (perhaps more often than not) they show varying degrees of phraseologization. The fact that these mappings are purely conventional — i.e., based on usage and the customary behavior of speakers — allows naturally for the development of specialized restricted uses of free phrases in certain pragmatic contexts (*pragmatemes*) or the establishment of set phrases as the "natural" way of expressing a particular complex meaning (*clichés*). The conventional nature of the meaning $\Leftrightarrow$ form relation also allows for the development of construction-specific meanings of particular forms in a particular phrasal context, giving rise to *collocations*.

Each of the aforementioned types of restricted phrase — the pragmateme, the cliché, and the collocation — falls under the heading of compositional phrasemes. A compositional phraseme is a restricted phrase AB whose meaning is a regular sum of the meanings of its constituents, but at least one of its components cannot be selected freely. The selection of the non-free constituents of a compositional phraseme can be restricted in one of three ways, giving us the three subtypes of compositional phraseme listed in Table 2, considered in descending order of restrictedness:

Pragmatemes: phrasemes wherein none of the components is selected freely and the restrictions are imposed by the situation of utterance — e.g., *Will you marry me?* [when making a marriage proposal] (cf. Rus. *Bud'(te) moej*

ženoj! 'Be my wife!'), *Best before* ... [on a container of packaged food] (cf. Rus. *Srok godnosti* – ... lit. 'Deadline of fitness is ...'), etc.

Clichés: phrasemes wherein none of the components is selected freely and the restrictions are imposed by conventional linguistic usage — e.g., *in the wrong place at the wrong time*; *You've seen one, you've seen 'em all!*; etc.

Collocations: phrasemes wherein one component of the phrase, the *base*, is selected freely, while the other, the *collocate*, is not, but is selected as a function of the other member of the pair — e.g., *black* COFFEE, *heavy* RAIN, *sound* ARGUMENT, BEAT [N] *to a pulp*, *pay* ATTENTION, *a* CLOUD *hangs* [*over* N], *fall into an* AMBUSH, CROWS *caw*, etc.[7]

Each of these phraseme-types differs in the source of the restrictions on the selection of one or more of its constituents, but all show some degree of paradigmatic restriction on their members. Collocations, in which one of the members (the base) is selected freely, are *semi-restricted* phrasemes, while pragmatemes and clichés, none of whose members are selected freely, are *fully restricted*.

The other property shared by all three types of phraseme is that they are semantically compositional. This is true not only in the obvious case of pragmatemes such as *Will you marry me?* and clichés such as *in the wrong place at the wrong time*, but also in the case of collocations such as *sound argument* or *heavy rain*. This follows from the definition of compositionality outlined above. Consider: if a complex sign AB of language ʟ that expresses the meaning 'S' is compositional, this means that 'S' can be divided into 'S$_1$' and 'S$_2$', such that 'S$_1$' corresponds to A and 'S$_2$' corresponds to B. In the case of a free phrase, the meanings of both A and B are their *inherent signifieds* — that is, the meanings that A and B express in the phrase correspond to *lexical units* (LUs) of ʟ (what are often referred to as "lexical entries"). If AB is a fully restricted phraseme (a *pragmateme* or a *cliché*), both A and B express their inherent signifieds and the meaning of AB is equal to the sum of A's meaning and B's meaning. Therefore AB is compositional in the same sense as any free phrase (although of course pragmatemes and clichés must be included in the lexicon along with the pragmatic or conventional restrictions on their usage).

The case of semi-restricted phrasemes (*collocations*), on the other hand, is somewhat more complicated. If AB is a collocation, then only the part that is selected freely, say A, necessarily expresses its own inherent signified, 'S$_1$'. A serves as the base of the collocation and the other member of the collocation — that is, B — expresses 'S$_2$'; however, 'S$_2$' is not always B's inherent signified: instead, B often takes on a construction-specific meaning, 'S$_2$''', this being B's *contextual* or *contingent signified*. Consider, for example, the English collocation *black coffee* 'coffee without addition of a dairy product'. Here, COFFEE

is A, the base, and it is selected freely according to the desired meaning of the Speaker. BLACK, which is B in this collocation, expresses the meaning 'without addition of a dairy product' only in combination with COFFEE. The signified 'S_2''' is therefore B's contingent signified, which occurs only within the collocation AB. The adjective B′ 'S_2''' — BLACK 'without addition of a dairy product' — is so restricted as not to be worth including in the English lexicon as a full-fledged LU. Instead, it is best described under COFFEE, as part of one of that lexeme's collocations. Lexemes often have construction-specific or contingent signifieds of varying degrees of semantic distance from what would be considered their inherent signified, a fact that leads certain linguists to treat compositionality as a gradient rather than a binary property (i.e., collocations where a contingent signified seems more closely related to the lexeme's inherent signified are often said to be more "compositional" than those in which the relation between signifieds is less transparent). We would maintain, however, that as long as the signified of the base of the collocation can be attributed to A as its inherent signified, the expression is completely compositional, whatever the semantic distance between the inherent signified of B and its contextual signified in the collocation.

Our example of BLACK COFFEE also serves to illustrate another key concept in the semantic representation of collocations and other restricted expressions — the *semantic pivot*. For an expression AB having the meaning 'S' such that 'S' = 'S_1' ⊕ 'S_2', the semantic pivot of 'AB' is that part 'S_1' of AB's meaning 'S' such that 'S' [= 'S_1' ⊕ 'S_2'] can be represented as a predicate 'S_2' bearing on 'S_1', i.e., 'S' = 'S_2'('S_1') (Mel'čuk 2006b: 277). In the case of BLACK COFFEE, the semantic pivot of the expression is COFFEE ('black coffee' = 'without the addition of a diary product'('coffee')).[8] This is true of all collocations, and represents an important distinction between collocations and noncompositional phrasemes (i.e., idioms), none of whose components express their semantic pivot. The concept of the semantic pivot will play an important role in the discussion below, particularly when it comes to describing the nature of morphological phrasemes in the realm of inflection.

2.2. *Noncompositional phrasemes*

A noncompositional phraseme is an expression none of whose members are selected freely (that is, they are fully restricted) and whose meaning is not compositional in the sense defined in Subsection 2.1 — in other words, whose meaning is not the regular sum of the meanings (inherent or contextual) of its parts. Noncompositional phrasemes are in fact idioms — e.g., ⸢POP THE QUESTION⸣ 'make a marriage proposal', ⸢CHEEK BY JOWL⸣ 'in close association', ⸢HAVE IT IN [FOR N_Y]⸣ 'intend to harm Y', ⸢HIT THE HAY⸣ 'go to bed', etc. These

expressions consist of words with clearly definable meanings, but the meaning of the idiom itself is not the sum of these, and none of the constituents can be said to be its base or to express the idiom's semantic pivot.

Like compositional phrasemes, noncompositional phrasemes can be divided into three major types, although in this case the division is based on to what degree (if at all) the meanings of the various constituents of the expression are included in meaning of the idiom:[9]

> *Weak idiom*: an idiom that includes the meaning of all its components, but none in the position of the semantic pivot — e.g., ⌜LIGHTNING ROD⌝ 'DEVICE that is designed to protect constructions against **lightning** and has the form of a conducting **rod**, fixed . . .' includes the meanings of both its constituents (boldfaced), but its semantic pivot (indicated by SMALL CAPS) is 'device'.
>
> *Semi-idiom*: an idiom that includes the meaning of only one of its constituents, but not in the position of the semantic pivot, while the meaning of the other is not included — e.g., ⌜SEA DOG⌝ 'MAN having significant experience of navigation on the **seas**' includes the meanings of one of its constituents, *sea*, but its semantic pivot is 'man'.
>
> *Strong idiom*: an idiom that does not include the meanings of either of its constituents — e.g., ⌜SHOOT THE BREEZE⌝ 'chat leisurely' does not contain the meanings 'shoot' or 'breeze'.

It is important to keep in mind a crucial feature of these definitions: the semantic pivot of an idiom is not expressed directly by any of the idiom's constituents. This distinguishes idioms from collocations, and will become a crucial point in the discussion of certain types of noncompositional morphological expressions in the sections below.

3.   Free and non-free complex morphological expressions

Issues of paradigmatic and syntagmatic freedom at the level of phrases are, of course, familiar and well-trodden ground. Even if the terminology and taxonomy introduced in the preceding sections are in some respects new, the underlying concept of the phraseologization of phrases is certainly not. The central point of this paper, however, is that phraseologization is a more general process that is not restricted to a particular type of sign: restrictedness and noncompositionality characterize not only phrases, but apply to all complex linguistic signs. Most importantly for us in this paper, they can and do apply below the syntactic (phrasal) level and characterize signs at the morphological level — at the level of the word. Words show the same properties of phraseologization as phrases, and (with some restrictions) phraseologized morphological expres-

sions, or *morphological phrasemes*, fall into the same major types and sub-types as phrasemes do on the syntactic level.

Although this parallel between phrasal and morphological signs has gone largely unnoticed, it should not come as a surprise when the morphological sign is considered from the perspective of the mapping of its semantic content to surface phonological form (that is, from the perspective of text synthesis). In the approach advocated in this paper, morphological expressions are treated on a par with all other types of linguistic sign, as conventionalized associations of meanings (signifieds) to forms (signifiers). As with phrases, the associations of morphological meanings and forms are potentially multivariate, in the sense of allowing one-to-many (synonymy) and many-to-one correspondences (poly-semy). They are also multistratal, in the sense that the mapping between the meaning of a morphological sign and its surface form is most conveniently modeled in terms of a series of intermediate abstract representations (Mel'čuk 1982, 1993–2000; Aronoff 1994). The most important intermediate level of representation for the purposes of this paper is the morphemic level. A *morpheme* is considered here to be a set of morphological elementary signs (*morphs*) that are conventionally associated with the same particular meaning and are contextually distributed according to general rules (i.e., rules that mention classes of, rather than individual, signs) based on their phonological and morphological properties and/or their environment.[10] Adding this intermediate level of representation provides at least two important advantages:

1) It side-steps the problems (discussed by Anderson [1992] and Stump [2001], among others) created by a simplistic "one-form, one-meaning" conception of morphology;

2) It allows direct consideration of the problems of mapping between the meaning of a morpheme and the corresponding set of signifiers (its morphs) without getting bogged down in considerations of the formal — that is, (morpho)phonological — relations between the concrete members of that set (for a discussion of this problem in slightly different terms, see Aronoff [1994: 22–23]).

The latter of these points is especially important here, as it is the conventionalized nature of this mapping between meaning and morpheme that lies at the heart of our discussion.

In the simplest case, the free complex morphological expression, there is a straightforward one-to-one mapping between a meaning and a morpheme, which is the set of signs having this meaning but different phonological forms. This mapping can be represented as a (Deep-)Morphological Rule such as that in (3), which models the expression of a single value of an inflectional category (a.k.a. a *grammeme*),[11] COMPLETIVE (a value of the category Aspect), that has as

its marker a single morpheme, {COMP} — the set of allomorphs of the completive marker:[12]

(3)   COMPLETIVE $\Leftrightarrow$ {COMP}

This is precisely parallel to the kind of rule that would be used to model the expression of a lexical meaning by an individual lexeme:

(4)   'fast$^2_1$' $\Leftrightarrow$ {FAST$^2_1$} (adjective; $\approx$ 'moving quickly')[13]

Just as a morpheme is a set ({...}) of morphs, so a lexeme is a set of lexes (word forms as well as phrases that express analytical inflectional forms).

   The rule in (3) models a simple, or minimal, morphological expression, while that in (4) models a simple, or minimal, lexical expression. Complex lexical meanings map onto complex (phrasal) expressions, which can be described by the combination of rules of the type shown in (4) into formulas like that in (5):

(5)   'very' $\oplus$ 'fast$^2_1$' $\Leftrightarrow$ {VERY} $\oplus$ {FAST$^2_1$}

The lexical expression on the right side of (5) is a free phrase (*very fast*), as defined in Section 2. It finds its parallel on the morphological level in free morphological expressions such as that illustrated in (6a), the Hungarian nominal form, *nyelvekben* 'in languages':

Hungarian
(6)   a.   *nyelvekben*
             nyelv–ek–ben
             language–PL–INESSIVE
             'in languages'
       b.   'language' $\oplus$ PLURAL $\oplus$ INESSIVE $\Leftrightarrow$ {NYELV} $\oplus$ {PL} $\oplus$ {INESSIVE}
       c.   'language' $\Leftrightarrow$ {NYELV}
             PLURAL $\Leftrightarrow$ {PL}
             INESSIVE $\Leftrightarrow$ {INESSIVE}

The noun in (6a) is inflected for plural number and the inessive case, while (6b) shows the mappings between its constituent morphemes on the right side of the formula (composed from the individual rules given in [6c]) and the meanings on the left that each individually expresses. Once again, this shows a compositional pairing of three freely chosen meanings with three morphemes used to express them — in short, it describes a free complex morphological expression.

   Like free phrases, restricted phrases (i.e., phrasal phrasemes) can also be modeled in terms of rules. Consider, for example, the strong idiom ⌐SHOOT THE BREEZE¬ illustrated in (7):

(7)   'chat' $\oplus$ 'leisurely' $\Leftrightarrow$ ⌐SHOOT THE BREEZE¬

In this case, none of the lexical elements on the right side of the rule are direct expressions of any particular semantic element on the left side. This is possible because of the conventionalized nature of meaning-to-form mappings, which allows speakers to move away from the simplistic principles of one-to-one correspondences and to assign new meanings to particular combinations of signs (or, in the case of collocations, contingent signifieds to signs in particular contexts). The resulting expressions are phrasemes. It is the central claim of this paper that, just as rules for free lexical expressions such as that in (5) find their parallels in rules for free morphological expressions like that in (6), rules for phrasemes such as the idiom in (7) find their parallel in morphological rules such as (8), which models the way in which the UNT first-person plural exclusive subject is expressed on the finite verb in the example in (1c) above:

(8) $1_{\text{SUB}} \oplus \text{PL}_{\text{SUB}} \oplus \text{EXCL} \Leftrightarrow \{1\text{SG}_{\text{SUB}}\} \oplus \{1\text{PL}_{\text{SUB}}:\text{INCL}\}$

This rule, like that in (7), describes an idiom: neither of the morphemes on the right side of the rule is a direct expression of any of the inflectional values on the left side. As we will argue in the following sections, morphological expressions show the same patterns in mappings between meanings and forms that are shown by phrases, including both free morphological expressions (Subsection 3.1) and morphological phrasemes (Subsection 3.2).

3.1.   *Free morphological expressions*

Free morphological expressions are words or complex parts of word forms whose components are chosen by the Speaker unrestrictedly and which are compositional. Such expressions are found in both derivation and inflection. The derivational type of free morphological expression in UNT is illustrated in (9), which presents some deverbal result nouns formed with the suffix *-mą* 'by-product of', added to a verbal stem with the signified 'perform action P upon entity α' in order to express the meaning 'substance that remains as the by-product of the action P by somebody upon entity α':[14]

(9) | | | | |
|---|---|---|---|
| *čį́tmą* 'bagasse' | < | *čįt-* 'press [= P] sugarcane [= α]' | *-mą* |
| *kiłwám* 'food scraps' | < | *kiłwá* 'nibble at, eat parts of some food' | *-mą* |
| *ląʔšʼá:m* 'wood or bamboo splinter' | < | *ląʔšʼá:* 'split wood or bamboo' | *-mą* |
| *pás̜mą* 'bare corn cob' | < | *pąs-* 'remove kernels from corn' | *-mą* |
| *pałštúm* 'floor-sweepings' | < | *pałštú* 'sweep out an area' | *-mą* |
| *šʼa:m* 'corn husk' | < | *šʼa:* 'husk corn' | *-mą* |

In (9), each of the two component parts of the derived noun is chosen (by the Speaker) freely for its meaning, the verbal radical P expressing the meaning 'action P upon entity α', while the suffix expresses the meaning 'substance

that remains as the by-product of the action P by somebody upon entity α'. The expression is paradigmatically unrestricted: P can be any semantically-compatible verb, and while the suffix *-mą* cannot be replaced, this is only due to the lack of another synonymous affix in the language (however, the same meanings could be conveyed by periphrastic expressions). The derived noun is thus a free derivative, unrestricted and compositional. Each morphological "piece" of it corresponds to a specific part of its semantic make-up, and its meaning is predictable from the inherent signifieds of its component parts.

The inflectional type of free morphological expression can be illustrated by second- and third-person object agreement of the UNT verb, shown in (10):

(10)　a.　*musú:*
　　　　　Ø–Ø–Ø–musú:–ya:
　　　　　3_{OBJ}–SG_{OBJ}–$3SG_{SUB}$–kiss–INCOMP
　　　　　's/he kisses him/her'

　　　b.　*ka:musú:*
　　　　　Ø–ka:–Ø–musú:–ya:
　　　　　3_{OBJ}–PL_{OBJ}–$3SG_{SUB}$–kiss–INCOMP
　　　　　's/he kisses them'

　　　c.　*musu:yá:n*
　　　　　Ø–Ø–musú:–ya:–n
　　　　　SG_{OBJ}–$3SG_{SUB}$–kiss–INCOMP–2_{OBJ}
　　　　　's/he kisses you$_{SG}$'

　　　d.　*ka:musu:yá:n*
　　　　　ka:–Ø–musú:–ya:–n
　　　　　PL_{OBJ}–$3SG_{SUB}$–kiss–INCOMP–2_{OBJ}
　　　　　's/he kisses you guys'

UNT marks the person and the number of the direct object with separate affixes, the third-person for objects being a zero prefix and the second-person being marked with the suffix *-n*.[15] Object number is marked by one of two prefixes, *Ø-* for 'singular object' and *ka:-* for 'plural object'. In all of the word forms in (10), the choice of object-person and object-number prefix is made freely, based on the meaning the Speaker wishes to express (specifically, on who and how many are being kissed), and the morphemes involved in these expressions can be handled descriptively by individual (Deep-)Morphological Rules, given in (11):

(11)　3_{OBJ}　⟺　$\{3_{OBJ}\}$
　　　　2_{OBJ}　⟺　$\{2_{OBJ}\}$
　　　　SG_{OBJ}　⟺　$\{SG_{OBJ}\}$
　　　　PL_{OBJ}　⟺　$\{PL_{OBJ}\}$

Each of these rules maps a single value of an inflectional category — that is, a grammeme — onto an individual morpheme, and none of the rules makes reference to any of the others. The forms in (10) are created by freely combining the application of one of the rules for the expression of object-person with that of a rule for the expression of object-number. In this sense, the forms in (10) are unrestricted. And, just as with the derivational free morphological expression in (9), each of the word forms in (10) is compositional, and its meaning is predictable from the inherent signifieds of its component parts.

3.2.   *Morphological phrasemes*

In contrast to free morphological expressions, where all of the constituent morphemes are selected unrestrictedly based on the intended meaning of the Speaker, morphological phrasemes are morphological expressions in which one or more of the constituent morphemes is selected in a restricted manner, depending on factors other than its specific meaning. This property of complex signs, paradigmatic freedom, is precisely the same as that used to distinguish free phrases from phrasal (i.e., lexical-syntactic) phrasemes in Section 2. The second property of complex signs discussed in Section 2, syntagmatic freedom, can also be applied to morphological expressions, allowing us to distinguish between compositional and noncompositional morphological phrasemes. Thus, up to this point the parallel between type of phrasal sign and type of morphological sign is exact; however, in the consideration of morphological phrasemes it is also necessary to take into account an additional axis of classification: the distinction between a *compound* or *derivational* morphological phraseme and a grammatical, or *inflectional*, morphological phraseme.[16] In the former case, the morphological phraseme would be of one of two types: 1) a phraseologized combination of several radical morphemes (a *compound morphological phraseme*, or 2) either a phraseologized combination of a radical morpheme plus derivational affixal morphemes (a *full derivational morphological phraseme*) or a phraseologized combination of two or more derivational affixes — an affix-complex — that combines with a variety of radicals (a *derivational affixal phraseme*).[17]

In the case of inflectional phrasemes there are fewer types to discuss, given that a compound inflectional morphological phraseme is improbable: the use of radical morphemes (let alone phraseologized combinations of radical morphemes) as inflection is not widespread.[18] However, the distinction between full and affixal inflectional phrasemes does apply in that it is possible to have both restricted stem + inflectional affix (a *full inflectional morphological phraseme*) and inflectional affix + inflectional affix combinations (an *inflectional morphological affixal phraseme*), although the former case is the only

possibility for compositional inflectional phrasemes. Noncompositional full-morphological phrasemes consisting of a combination of a stem morpheme + inflectional affixal morpheme(s) are ruled out by the nature of inflection. Simply put, if such a combination of elements were to have an idiomatic meaning, this would constitute the phraseologization of a whole inflectional form for derivational purposes, and thus it would no longer be an inflectional form. Such cases exist — for instance, in Spanish where the plural inflection of the lexeme PADRE 'father' has given rise to another lexeme PADRES 'parents' (homophonous with *padres* 'fathers'), or in Upper Necaxa Totonac, where the progressive form of the verb NIː 'die', *niːmáːɬ*, has an idiomatic meaning 'be sick', which it does not have in any other aspectual form. In both cases, the homophonous word form in the appropriate inflection — that is, a free morphological expression (Sp. *padres* 'fathers' and UNT *niːmáːɬ* 's/he is dying') — coexists with the phraseologized expression (Sp. PADRES 'parents' and UNT NIːMAːɬ 'be sick'). The same reasoning applies to the combination of derivational and inflectional morphemes: if the combination has a phraseologized meaning, this is necessarily a derivational use of the inflectional elements. Thus, inflectional morphological phrasemes break down into only three categories: compositional full inflectional phrasemes, compositional inflectional affixal phrasemes, and noncompositional inflectional affixal phrasemes.

In the remainder of this paper we will examine in more detail and exemplify each of the types of morphological phraseme proposed here. Treating derivational and compound morphological phrasemes together for the purposes of exposition allows us to break the discussion down into four broad categories — compositional derivational phrasemes (subdivided between full and affixal), compositional inflectional phrasemes (subdivided also between full and affixal), noncompositional derivational phrasemes (likewise subdivided between full and affixal), and noncompositional inflectional phrasemes. We will deal with each of these types in turn.

3.2.1. *Compositional compound and derivational phrasemes.* At the phrasal level, compositional phrasemes can be divided into three types based on the source of the restrictions governing the selection of their non-free member: pragmatemes (in which the selection of the constituent elements is restricted by the pragmatic situation), clichés (in which the selection of all the constituent elements is restricted by linguistic convention), and collocations (in which the selection of one of the constituent lexemes is restricted by the other member of the expression). At the level of morphology, however, it seems that only the third of these types, the morphological collocation, is well-attested. The absence of the other two logical possibilities, the morphological pragmateme and the morphological cliché, may be a consequence of the nature of these two types of phraseme: both would necessitate that some composi-

tional combination of morphemes be required (by pragmatics or by convention) in some context where there exists one or more completely synonymous combinations of other morphemes that are not allowed, in spite of a perfect semantic fit. Given the numerically-limited nature of the morphological resources of a language (numbering at most a few hundred, as opposed to the lexical resources, which are numbered in the hundreds of thousands, a difference of three orders of magnitude), the existence of completely synonymous combinations of morphemes is in itself unusual, and so it seems highly improbable that restricted selection among possible alternative morphological expressions of this sort would occur with any frequency. Still, we do not want to completely rule out the possibility that such things exist; we can only report that we have as yet failed to find any convincing examples thereof.

What are more robustly attested in natural languages are compound derivational collocations. From Russian we have examples such as *kon–e–vodstv-(o)* 'horse breeding' vs. **lošad–e–vodstv-(o)*, (KON′ 'horse' and LOŠAD′ 'horse' being approximate synonyms), in which the selection of the stem meaning 'horse' is determined by the nominal stem *vodstv-* 'breeding' (which is the base of the expression). Similarly, *xleb* 'bread' is permitted in combination with -*rob* 'grower' in *xleb–o–rob* 'bread grower' whereas *zern-(o)* 'grain' is ruled out with -*rob* (**zern–o–rob* 'grain grower'), although apparently not for semantic reasons — viz. *zern–ov-(oe) xozjajstvo* 'grain-growing farm'. In English, we have compounds such as *taxi stand* (cf. *cab stand), *soccer pitch* (*baseball pitch), and *burger joint* (cf. *pizza joint* but *French-fry joint, *falafel joint). In each of these cases, both members of the compound express their ordinary, inherent meanings, and so the compounds are compositional — yet certain analogous, semantically-plausible combinations are ruled out. Thus, *stand* 'place where vehicles wait for passengers' combines with *taxi*, but not with *cab*, to form an expression meaning 'place where taxicabs wait for passengers'; *pitch* 'field where a sport is played' combines happily with *soccer* to mean 'field where soccer is played' and not with *baseball* (which requires *diamond* to form the analogous compound). The slang term *joint* 'establishment for eating, drinking, or entertainment' can be used in combination with *burger* and *pizza*, but not with *French-fry* or *falafel*. And so on. For each of these compositional compounds, one of the components restricts the selection of the other in a manner exactly parallel to the collocational restrictions seen in phrasal collocations and in compositional derivational phrasemes. Unfortunately, this type of compounding is not productive enough in Upper Necaxa for us to provide examples from that language.

3.2.1.1.   *Full derivational collocations.*   Full derivational collocations are compositional morphological expressions that are restricted in that not all of their component morphemes are chosen freely. Instead, one of the morphemes

is chosen as a function of the other morphological constituent of the expression, its base. This type of situation is quite familiar in derivation, as selectional restrictions placed by radicals on (near-)synonymous derivational affixes are well-known (e.g., Fabb 1988; Giegrich 1999; Plag 2002). Although we have no clear examples from UNT, two cases from English would be the choice of nominalizer as conditioned by particular verbal bases (e.g., *payment*, **payation*; *suffixation*, **suffixment*; etc.), and the selection of inhabitant suffixes for particular place names (*Montrealer*, **Montrealian*; *Edmontonian*, **Edmontoner*; etc.); in both cases, the choice of derivational affixal morpheme is restricted by the individual radical (the base of the collocation). From Russian, we have examples of the selection of agentive suffix: *vodi-*(*t'*) 'drive' ~ *vodi–tel'* 'driver', but *čisti-*(*t'*) 'clean' ~ *čisti–l'ščik* 'cleaner' and *pečat-*(*at'*) 'print' ~ *pečat–nik* 'printer [typographical worker]'. The choice of suffix here is determined by the verb,[19] which is the base of the collocation. Nevertheless, the derived form remains compositional: each morpheme can be attributed a particular sub-part of the meaning of the whole, and each morpheme expresses an inherent meaning with which it is regularly associated in a wide range of other expressions in the language.

3.2.1.2.   *Derivational affixal collocations.*   Combinations of derivational affixes, one of which is chosen freely based on its meaning and the other of which is added automatically as its collocate, are somewhat harder to come by, although UNT offers a fairly clear example in the formation of its causative verbs. In Upper Necaxa, the causative takes one of two forms, depending on the inflection class (stative or dynamic) of the base verb:

(12)   a.   *ą̓ʔé:* 'be uncovered'   >   *mą:ʔé:* 'uncover something'
            ʔe:nú: 'be off to one side'   >   *ma:ʔe:nų́* 'put something aside'
            tsumá: 'be full'   >   *ma:tsumá:* 'fill something'
       b.   *teŋwán* 'get damp'   >   *ma:teŋwaní:* 'dampen something'
            łi: 'dance'   >   *ma:łi:ní:* 'make someone dance'
            pą̓ʔɬ- 'break something'   >   *ma:pą̓ʔɬní:* 'make someone break something'

When the verb is stative (12a), its causative (with a few exceptions) is formed with the causative prefix *ma:-* alone; dynamic verbs (12b) form their causative by the combination of *ma:-* and a suffix *-ni:*, which historically was a transitivizer but is now restricted to appearing in causatives and a few fossilized forms (Beck 2004). In the causative of dynamic verbs, the prefix is chosen based on its meaning and the suffix is an automatic collocate. Because *ma:-* has the inherent meaning 'cause', the semantic contribution of the suffix to the construction is nil, and the affix-complex as a whole is thus compositional in the sense that the term is being used in this paper.

3.2.2. *Compositional inflectional phrasemes.* Perhaps more novel than derivational and compound morphological collocations are inflectional morphological collocations — that is, morphological expressions of inflectional meanings such that at least some of their elements are restricted rather than free. As with compositional compound and derivational phrasemes, two of the potential subtypes of compositional phraseme, the pragmateme and the cliché, are unattested: to date, we have identified only inflectional collocations. Of these there are two types — the *full* inflectional collocation, discussed in Subsection 3.2.2.1, and the affixal inflectional collocation or parasitic formation, discussed in Subsection 3.2.2.2.

3.2.2.1. *Full inflectional collocations.* A full inflectional collocation is a phraseologized combination of a stem morpheme and an inflectional morpheme such that this particular inflection is chosen restrictedly based on this individual stem. An example of a full inflectional collocation from a familiar language can be found in the expressions of Russian verbal aspect. For many verbs, this distinction is marked by using various prefixal morphemes such as those shown in Table 3 to indicate perfective aspect. All in all, there are 15 perfective morphemes. Semantically, the perfective forms are compositional: all the prefixes express the perfective aspect. However, their selection cannot be described in general or systematic terms (hence, the claim that there are 15 perfective morphemes rather than 15 suppletive allomorphs of a single perfective morpheme): each verb has to be marked in the lexicon for the particular perfective prefix morpheme it takes. The choice of perfective prefix for each radical is thus severely restricted, but the resulting word is compositional. This is characteristic of collocations, and exactly parallels the phrasal and derivational collocations described above. The stem is the base of the collocation, as well as its semantic pivot.

   Another example of an inflectional morphological collocation is the plural form of nouns in Burushaski. This language has about 70 plural suffix morphemes, some of which are shown in Table 4, distributed in an unpredictable way. The plural forms of nouns are semantically compositional (consisting of

Table 3.   *Russian verbal aspect prefixes (Švedova 1980: 587–588)*

	IMPERFECTIVE	~	PERFECTIVE	
'read'	*čitat'*	~	*pro-*	*čitat'*
'try to wake'	*budit'*	~	*raz-*	*budit'*
'do'	*delat'*	~	*s-*	*delat'*
'build'	*stroit'*	~	*po-*	*stroit'*
'drink'	*pit'*	~	*vy-*	*pit'*
'make drink'	*poit'*	~	*na-*	*poit'*

Table 4. *Burushaski nominal plural suffixes (Berger 1974: 15–20)*

	SINGULAR	PLURAL	
'king'	*thám*	*thám*	*-u*
'bread'	*páqu*	*páqu*	*-mu*
'dragon'	*aiždahár*	*aiždahár*	*-išu*
'branch'	*táγ*	*taγ*	*-ášku, -šku*
'pigeon'	*tál*	*tál*	*-ǯu*
'stone'	*dán*	*dan*	*-ǯó*
'enemy'	*dušmán*	*dušmá*	*-yu*
'rock$_N$'	*čár*	*čar*	*-kó*
'dog'	*húk*	*huk*	*-á, -ái*
'wolf'	*úrk*	*urk*	*-á, -ás*
'man'	*hír*	*hur*	*-í*
'demon'	*díu*	*diw*	*-ánc*
'flower'	*asqór*	*asqór*	*-iŋ*
'plow'	*hárč*	*harč*	*-óŋ*
'wind'	*tiṣ*	*tiṣ*	*-míŋ*
'minister'	*waziir*	*waziir*	*-ting*
'woman'	*gús*	*guš*	*-íngants*
'[a] mute'	*gót*	*goṭ*	*-ó*
'body'	*ḍím*	*ḍím*	*-a*
'horn'	*túr*	*tur*	*-iáŋ*
'saber'	*gaté+nč*	*gaté*	*-h*
'walnut'	*tilí*	*tilí*	——

a radical morpheme expressing the lexical meaning and a suffixal morpheme expressing PLURAL), but for each individual radical (the base of the collocation and its semantic pivot), the corresponding plural suffix has to be learned. The distribution of the different plural morphemes is completely idiosyncratic and does not correspond to any more general morphological or declension class of nouns in the language: therefore, it cannot be described as allomorphy, which should be sufficiently general as to be treated as rule-governed behavior.[20]

3.2.2.2. *Affixal inflectional collocations: Parasitic formations.* Another, rather different, type of morphological collocation is the so-called *parasitic formation* (Matthews 1972), which can be defined as follows:

> The expression of a particular value of an inflectional category (grammeme) G by the morpheme {G} is called a *parasitic formation* if and only if the use of {G} obligatorily entails adding the morpheme {G$_1$} to the stem, {G$_1$} being the inherent expression of another inflectional value G$_1$, but in this combination being empty.

In other words, in a parasitic formation we have both

$G \Leftrightarrow \{G\} \oplus \{G_1\}$

and

$G_1 \Leftrightarrow \{G_1$

In parasitic formations, the morpheme $\{G\}$ is chosen freely; it is the base of the expression, while the selection of $\{G_1\}$ is restricted — being conditioned by the selection of $\{G\}$. Such expressions are compositional in the technical sense, in that the meaning G can be attributed to $\{G\}$, while a contingent empty meaning can be attributed to $\{G_1\}$ (i.e., $\{G_1\}$ has no meaning in this construction — cf. the use of semantically empty "light" verbs in expressions such as *do a favor* or *make a mistake*). However, unlike the base of phrasal and derivational collocations, which must express the semantic pivot of the collocation, the base of the parasitic formation is not its semantic pivot — that is, $\{G_1\}$ is not the expression of a semantic predicate bearing on G. Rather, the two morphemes, taken together, form an inflectional affix-complex that, in an actual word form, bears semantically on the stem to which they are attached. Inflectional affix-complexes in and of themselves never form independent expressions, nor are they uniquely associated with particular lexical units. Thus, they cannot be said to have semantic pivots at all, making the parasitic formation a *pivotless collocation*. As we shall see in Section 3.2.4, the lack of a semantic pivot is typical for inflectional phrasemes of other types as well.

A well-known example of a parasitic formation can be found in the case-paradigms of many Daghestanian languages (Mel'čuk 2006a: 457–459). Archi, for example, inflects its nouns for 24 cases, four of which are illustrated in Table 5 for the noun GEL 'cup'. Like other languages in its family, the regular declension of Archi nouns forms its oblique cases other than the ergative (i.e., genitive, dative, etc.) based on the ergative case form of the noun. Thus, the ergative singular form of GEL is *gélli*, and all the remaining singular forms in the oblique cases are based on *gélli* (rather than on the radical *gél*). Likewise, in the plural the oblique case forms are based on the ergative plural, *gélumčaj*, rather than on the nominative plural *gélum*.[21] This can be illustrated by the following rules:

Table 5. *Some Archi case forms for* GEL *'cup' (Kibrik 1997: 27–28)*

Case	Number				
	SINGULAR		PLURAL		
NOMINATIVE	*gél*		*gél*	*-um*	
ERGATIVE	*gél*	*-li*	*gél*	*-um*	*-čaj*
GENITIVE	*gél*	*-li*	*-n*	*gél*	*-um*	*-če*	*-n*
DATIVE	*gél*	*-li*	*-s*	*gél*	*-um*	*-če*	*-s*

(13)   NOM ⟺ {NOM}
     ERG ⟺ {ERG}
     GEN ⟺ {ERG} ⊕ {GEN}
     DAT ⟺ {ERG} ⊕ {DAT}

The oblique case forms themselves do not express the ergative case, although they contain the ergative suffix (Mel'čuk 2008). Thus, the suffix-complexes *-li-n*, *-li-s*, etc., are inflectional collocations. The oblique case suffixes (i.e., *-n*, *-s*, etc.) are the bases, and the empty suffixes *-li* and *-če* are collocates: they are automatically added in order for the suffix-complex to be well formed. Because the empty suffixes are meaningless inside a parasitic formation, they cannot be described as the expressions of semantic predicates; therefore, the base of the formation — the full suffix — is not its semantic pivot. Rather, in the word form *géllin* 'cup$_{SG, DAT}$' the entire affix-complex *-li-n* bears on the stem *gél*, making *gél* the semantic pivot of the whole word form, and the parasitic formation itself a pivotless collocation.

    An example of a parasitic formation in UNT can be found in the verbal aspectual paradigm — specifically, in the expression of the progressive aspect,[22] whose marker is a combination of the suffixes *-ma:* 'progressive' and *-li̯* 'completive':

(14) a. *pasá:ł*
         pasá:–li̯
         burn–COMP
         'it burns$_{COMP}$ up'
    b. *pasa:ʔó:ł*
         pasá:–ʔo:–li̯
         burn–TOT–COMP
         'it burns$_{COMP}$ up completely'
    c. *pasa:má:ł*
         pasá:–ma:–li̯
         burn–PROG–COMP
         'it is burning' [= PASÁ:$_{PROG}$]
    d. *pasa:ma:ʔó:ł*
         pasá:–ma:–ʔo:–li̯
         burn–PROG–TOT–COMP
         'it is burning completely' [≈ 'it is completely on fire';
         = PASÁ:$_{PROG, TOT}$]

As shown in (14a), the completive on its own is expressed by a suffix *-li̯* (reduced here by morphophonemic processes to /ł/). The completive suffix follows all other suffixes, including the quasi-inflectional totalitative *-ʔo:* (14b). The progressive aspect, on the other hand, is expressed by the affix-complex

-ma: ⊕ *-li̯*, shown in (14c). The meaning of the expression is that the object in question is on fire or in the process of burning; it does not mean 'it is burning up', in spite of the presence of *-li̯*. The progressive suffix itself (that is, the suffix uniquely associated with the progressive aspect), *-ma:*, appears closer to the stem than the completive suffix; other suffixes — like the totalitative in (14d) — can intervene between the two, showing clearly that they are separate affixes.[23] The realization of the progressive aspect thus requires a rule like that in (15):

(15)   PROGRESSIVE ⟺ {PROG} ⊕ {COMP}

This rule creates a suffix-complex that is a collocation in the sense defined in Subsection 2.2: the morpheme {PROG} is freely chosen, the base of the collocation, and the morpheme {COMP} is its automatic collocate. Neither of them bear semantically on the other, so that neither express the semantic pivot of the collocation. Furthermore, given that the morpheme {PROG} is never used without {COMP}, while {COMP} itself is a regular means of expressing the completive aspect, progressive aspect forms can be said to be typical parasitic formations.

   Another example of a parasitic formation can be found in the expression of the first-person plural object. As shown in Subsection 3.1, UNT regularly expresses the person and number of objects with separate affixes; however, the first-person plural object is expressed by an affix-complex, *kin-* ⊕ *ka:-* ⊕ *-n*. Consider the following forms:

(16)   a.   *kintú̯ksa*
             kin–Ø–Ø–tu̯ks–ya:
             1_{OBJ}–SG_{OBJ}–$3SG_{SUB}$–hit–INCOMP
             'he hits me'

        b.   *tu̯ksá:n*
             Ø–Ø–tu̯ks–ya:–n
             SG_{OBJ}–$3SG_{SUB}$–hit–INCOMP–2_{OBJ}
             'he hits you$_{SG}$'

        c.   *ka:tu̯ksá:n*
             ka:–Ø–tu̯ks–ya:–n
             PL_{OBJ}–$3SG_{SUB}$–hit–INCOMP–2_{OBJ}
             'he hits you$_{PL}$'

        d.   *kinka:tu̯ksá:n*
             kin–ka:–Ø–tu̯ks–ya:–n
             1_{OBJ}–PL_{OBJ}–$3SG_{SUB}$–hit–INCOMP–2_{OBJ}
             'he hits us' [= TU̯KS$_{INCOMP, 3SG.SUB, 1PL.OBJ}$]

(16a) shows the first-person object marked with the prefix *kin-* and (16b) the second-person object marked with the suffix *-n*. In (16c), the plurality of the

second-person object is marked by a separate prefix, the object-pluralizer *ka:-*. The selection of these affixes is completely free, the person and number markers being chosen based on the meaning the Speaker wishes to express. These combinations of morphemes are also compositional. Each value of the person and number categories of the object has its own marker, and the choice of markers is not specially restricted, as shown by the rules in (11) above (plus an additional rule given in (17a) below).

In (16d), on the other hand, the affix-complex expressing object agreement is not unrestricted. The exclusive/inclusive distinction is not maintained for objects, and so even if the first-person object prefix *kin-* and the plural object prefix *ka:-* combine compositionally, the second-person object prefix *-n* is added without any semantic justification — at least in those cases where the second-person is not affected by the event. This requires both the rule for the expression of the first-person object formulated in (17a) plus a special rule for the implementation of the plural object marker, given in (17b):

(17)   a. $1_{OBJ} \Leftrightarrow \{1_{OBJ}\}$
       b. $PL_{OBJ} \Leftrightarrow \{PL_{OBJ}\} \oplus \{2_{OBJ}\} \mid person_{OBJ} = 1$

According to (17b), PL_{OBJ} is expressed as the combination of the morphemes $\{PL_{OBJ}\}$ and $\{2_{OBJ}\}$ under the condition (specified in the rule to the right of the vertical line) that the object be first-person.[24] As with the rule for the expression of the progressive aspect in (15) above, the rule here creates a parasitic formation: when pluralizing a first-person object, the plural object prefix *ka:-* requires an empty second-person object suffix.

3.2.3.   *Noncompositional compound and derivational phrasemes: Compound and derivational idioms.*   A noncompositional phraseme, or idiom, is an expression none of whose elements are chosen freely and whose meaning is not compositional in the sense defined in Subsection 2.1 — in other words, whose meaning is not the regular sum of the (inherent or contextual) meanings of its parts. Furthermore, in an idiom the semantic pivot is never expressed by any of the idiom's components. On the level of word-formation, morphological idioms of all types are quite familiar and are quite widespread in the languages of the world (Haspelmath 2002: 73–75; Aikhenvald 2007), although they are seldom referred to as idioms,[25] being more commonly lumped together under the loose heading of "lexicalized forms." Like phrasal idioms, compound and derivational idioms can be divided into the three familiar types defined above — weak idioms (Subsection 3.2.3.1), semi-idioms (3.2.3.2), and strong idioms (3.2.3.3).

3.2.3.1.   *Weak compound and derivational idioms.*   In derivation and compounding, a weak morphological idiom is a noncompositional restrictedly de-

rived or compounded lexeme whose meaning includes the meanings of all of its constituents as well as some additional meaning or meanings that pertain to the lexeme as a whole, one of these additional meanings being its semantic pivot. Some examples of weak compound idioms in UNT are given in (18):[26]

(18) *kịwi:ʔọ:lụ́:* '**forest-dwelling** humanoid MONSTER that resembles an **old man**'

    < *kị́wị* 'forest-dwelling'   *ʔọ:lụ́:* 'old man'

šwaya:skị́:tị '**fish-eating** BIRD [= kingfisher]'

    < *šwáyạ* 'eater-of'   *skị́:tị* 'fish'

lakaxú:kị 'ANT with a **face** shaped like that of a **deer**'

    < *lakán* 'face'   *xú:kị* 'deer'

The first, *kị́wi:ʔọ:lụ́:,* is a compound noun composed of the nouns *ʔọ:lụ́:* 'old man' and *kị́wị*, which on its own means 'tree/wood' but which in compounds regularly takes on the contingent signified 'forest-dwelling' (cf. *kịwi:pášnị* 'boar' [lit. 'forest-dwelling pig (*pášnị*)'], *kịwi:čičị́* 'coyote' [lit. 'forest-dwelling dog (*čičị́*)'], etc.). The meaning of the compound as a whole is not precisely 'forest-dwelling old man', but rather is a descriptive expression for a type of monster, 'monster' being the semantic pivot of the expression. Similarly, *šwaya:skị́:tị* ('one who eats fish') is the name of a type of bird, but 'bird' — its semantic pivot — is not part of the meanings of either of its constituent morphemes. The same is true of *lakaxú:kị*, which is the name of a type of ant but which is composed of words meaning 'face' and 'deer', neither of which include the meaning of the semantic pivot of the expression, 'ant', in their own meaning. Thus, both of these latter forms are also examples of compound weak morphological idioms.

    Weak full derivational idioms are noncompositional combinations of a root morpheme with a derivational affix whose meaning includes the meanings of both of its components. The UNT examples in (19) fall into this category:

(19) *ạʔłtama:nạ́* 'WASP that builds a nest by **cementing piles** of mud to a surface'

    < *ạʔłtamá:* 'cement into pile'   *-nạ́* 'agentive'

loŋót 'DISEASE which causes person to shiver as if they **feel cold** [= malaria]'

    < *loŋ-* 'feel cold'   *-ot* 'result'

ạʔpíščị 'TRAP made from a noose and a bent sapling to catch prey **tying it by the neck**'

    < *ạʔpiščí* 'tie by neck'   *-ʔ* 'nominalizer'

The first example, *ạʔłtama:nạ́*, is based on the verb *ạʔłtamá:* 'cement into pile', combined with the agentive nominalizer *-nạ́*. It would have a literal meaning 'one who cements things into a pile', but instead it is the name of a

particular species of wasp. The minimal paraphrase of this expression would be 'wasp', and so 'wasp' is the semantic pivot of the expression, although it is not a part of the meaning of any one of the expression's constituents. Similarly, the semantic pivot of the noun *loŋót* 'malaria' is 'disease', although formally it is the combination of the verb *loŋ-* 'feel cold' and the resultative nominalizer, which (among other things) creates nouns that express a result, substance, action, or sensation which is definitive of the state/process expressed by the verbal radical. Neither of the component morphemes has the meaning 'disease' as part of its semantic make-up, although together they could plausibly express 'shivering' (that is, typical behavior of one who feels cold, or has malaria). The final example breaks down along similar lines: *aʔpíščị* is a type of trap, but 'trap' is not a part of the meaning of either of its component morphemes, *aʔpiščí* 'tie by neck' or the nominalizer (-ʔ being realized as leftward movement of stress and final laryngealization as a result of morphophonological processes — Beck 2004).

Weak derivational affixal idioms consist of derivational affix-complexes whose meaning includes the meaning of both its components, and in addition some added meaning that includes the complex's semantic pivot. UNT has an example of this type, a circumfix based on the combination of the body part prefix *pu:-* 'container' and the nominalizing suffix *-nị* 'thing involved in or affected by action P':

(20)   *pu:čapán* 'mill, grinder'   <   *čapá* 'grind something'
       pu:čịtnị 'sugarcane mill'   <   *čịt* 'mill sugarcane'
       pu:skuyú:n 'smoking rack'   <   *skuyú:* 'smoke something'
       pu:ma:tsinkí:n 'scale'   <   *ma:tsinkí:* 'weigh something'

Considered as a whole, the affix-complex *pu:-* X *-nị* has the meaning 'container-like INSTRUMENT used to perform action P'. This meaning includes both the meaning of the prefix and of the suffix; however, the semantic pivot of the affix-complex as a whole, 'INSTRUMENT', is not expressed by either of the component affixes. Thus, the *pu:-* X *-nị* circumfix conforms to the definition of a weak idiom.

3.2.3.2.   *Compound and derivational semi-idioms.*   In contrast to weak idioms, a semi-idiom includes the meaning of only one of its constituents, but not in the position of the semantic pivot, while the meaning of the other is not included. English compound semi-idioms are quite common, and include widely cited examples such as *pancake* 'FOOD made from batter cooked in a pan', *lumberjack* 'LABOURER who fells trees for lumber', and *gravyboat* 'elongated DISH with a spout for pouring gravy'. In UNT, on the other hand, compound semi-idioms are only sparsely represented. The nouns in (21) are good examples:

(21) *a̰ʔlta:pa:tasíuм* 'ʙᴇʟᴛ worn around the **waist**'
      < *a̰ʔlta:pá:n* 'waist'   *tasíuм* 'fiber'
    ʔe:ʔa̰:š 'ꜱʜᴇʟʟ on the **back** of an animal (e.g., turtle, armadillo)'
      < *ʔe:n* 'back'   *ʔa̰:š* 'gourd'

The first expression, used specifically for the belt worn with pants (which is different from the belt used for slinging a machete over one's shoulder, fixing a saddle to a horse, etc.) includes as part of its meaning the signified of the prefix *a̰ʔlta:pa:-* 'waist', but does not directly include the meaning of *tasíuм* 'fiber' (although there is, historically, a metonymic link, as belts were made from tree-bark fibers before the advent of leather). As the semantic pivot of the expression is 'belt', which is the signified of neither of the constituent morphemes, *a̰ʔlta:pa:tasíuм* meets all the requirements of a compound semi-idiom. Likewise, the semantic pivot of *ʔe:ʔa̰:š* is 'shell', which is referred to only figuratively by its component *ʔa̰:š* 'gourd', and so is not expressed directly by either of the compound's components. On the other hand, the non-pivot *ʔe:n* 'back' is a part of the meaning of the compound, making the expression a semi-idiom.

Full derivational semi-idioms — that is, derivational semi-idioms consisting of a root and one or more derivational affixes are also widespread in languages like English. As a stock example, we can cite the English noun *locker* 'usually metallic ᴄᴏᴍᴘᴀʀᴛᴍᴇɴᴛ that can be **locked**, designed for the safekeeping of clothing and valuables of an individual in a public place': it includes the meaning of the stem, but not as a semantic pivot. The meaning of the agentive suffix is not present in the meaning of the expression as a whole. The same can be said of other agentive nominals in English such as *runner* 'ꜱʜᴏᴇ worn while **running**' and *mixer* 'a ᴘᴀʀᴛʏ where single people **mix** socially'. Curiously, full derivational semi-idioms seem to be rare in UNT (or at any rate, none have turned up in our lexical database to date).

Derivational affixal semi-idioms are formed by phraseologized combinations of derivational affixes and include the meaning of one of those affixes but not the other. An example would be the Russian circumfix *za-*X-*j-* 'region located behind natural obstacle X':

(22) *Zabajkal'j-(e)* 'region east of [≈ beyond] Lake Baikal'
      < *za-* 'beyond'   *bajkal* 'Lake Baikal'   *-j*
    Zakarpat'j-(e) 'region east of [≈ beyond] the Carpathian mountains'
      < *za-* 'beyond'   *karpat-(y)* 'Carpathians'   *-j*
    Zavolž'j-(e) 'region east of [≈ beyond] the Volga river'
      < *za-* 'beyond'   *volg-(a)* 'Volga'   *-j*

The prefixal portion of the circumfix, *za-*, expresses its inherent meaning 'beyond'; however, the suffix *-j* is a bit harder to assign a meaning to. It is used in

other formations without *za-* where it has a variety of meanings (perhaps most consistently something like 'collective'), none of which seem to form a part of the meaning of this circumfix, suggesting that the suffix makes no semantic contribution to the phraseme at all. Under this analysis, the circumfix as a whole is noncompositional ('beyond' ⊕ 'Ø' ≠ 'REGION located beyond natural obstacle X'), making it an idiom. The inclusion of the meaning of the other component, *za-*, in the meaning of the expression as a whole makes this a semi-idiom.

Admittedly, an alternative would be to posit a contextual meaning of 'region around a natural obstacle X' to -*j* which is specific to this construction, in which case the expression becomes compositional in the sense that the meaning of the whole is the sum of the meanings of its parts. Aside from being rather ad hoc and implausible (the other uses of -*j* do not suggest a plausible diachronic path or semantic motivation for this meaning to have become uniquely associated with this suffix), this would make -*j* the semantic pivot and the *za-X-j-* circumfix a collocation in which the base (the expression of the semantic pivot) has a contextual meaning associated only with that collocation. By the same token, positing a contextual meaning for -*j* of 'region' might be used to argue that the circumfix, although noncompositional ('beyond' ⊕ 'region' ≠ 'REGION located beyond natural obstacle X') is a weak, rather than a semi-idiom — but again, this seems problematic from the semantic and diachronic points of view, and it is not clear what the advantages of this would be. The disadvantage would be that 'region' is the semantic pivot of the whole expression, which violates the definition of the idiom (whose components should never express the idiom's semantic pivot). Of course, the fact that a particular datum might counter-exemplify a proposed taxonomy is never an argument for dismissing that datum, but in this case there seem to be no compelling reasons to pursue either of the problematic interpretations. What is important about this discussion, however, is that it underscores the inherent difficulty that arises when dealing with phraseologized uses of very abstract or semantically bleached elements (be they bound morphemes or function words such as light verbs and auxiliaries): it is often very difficult to make air-tight cases for such elements having well-defined inherent meanings, and therefore it is often difficult to tease out their exact semantic contribution to phraseologized expressions of which they form a part.

3.2.3.3.   *Strong compound and derivational idioms.*   The third type of idiom is the strong idiom — that is, a noncompositional expression in which the meanings of none of the constituent elements are part of the meaning of the expression as a whole. Strong compound idioms are quite familiar, including as they correspond to what are sometimes referred to as "exocentric compounds" (Kiparsky 1982). In a strong compound idiom, the meanings of none

of the constituent stems are part of the meaning of the compound itself. A few UNT examples are given in (23):

(23)   *ta:ʔeštáx* 'Milky Way'
    <   *tex* 'path'   *ʔeštáx* 'lime'
    išli:što̱ʔotą́ʔo̱ 'Mullein nightshade'
    <   *iš-* '3po'   *li:štǫ́ʔ* 'needle'   *tą́ʔo̱* 'old woman'
    pi̱ni:kúču̱ 'ginger'
    <   *pi̱n* 'chili'   *-i:* 'connective'   *kúču̱* 'medicine'

The first of these compounds is the name of an astronomical object, but its components *tex* 'path' and *ʔeštáx* 'lime (calcium hydroxide)', are derived from the figurative description of the Milky Way, which looks like a path drawn in lime (a fine white powder) across the sky. The name of the spiny plant Mullein nightshade, *išli:što̱ʔotą́ʔo̱* (literally 'the old woman's needle'), is likewise a figurative expression, while the name given to ginger (literally 'chili-pepper medicine') is even more opaque. It is perhaps a reference to the spicy taste of the root, but this plant (an introduced species) is not particularly sought after as medicinal in Totonac culture. In none of these cases is the meaning of any of the words composing the compounds part of the meaning of the expression as a whole, putting them solidly into the camp of strong idioms.

   Strong full derivational idioms are combinations of a root and a derivational affix that together express a meaning that includes the meanings of neither component. These are quite common in UNT. (24) presents a few such expressions from the realm of nominalized verbs:

(24)   *kúni̱* 'caterpillar'   <   *kun* 'swell'   *-ni̱*
    skúxni̱ 'official'   <   *skux-* 'work'   *-ni̱*
    la̱ʔspútni 'dead person'   <   *la̱ʔspút-* 'be used up'   *-ni̱*
    pášni̱ 'pig'   <   *paš-* 'bathe'   *-ni̱*

Each of these morphological idioms is formed by the combination of a verb stem with the nominalizer *-ni̱*, whose usual function is to form a deverbal expression denoting something that is undergoing or has undergone the process expressed by the verb stem (*nomen patientis*). Thus, the expected meaning of *kúni̱* is 'swollen (one)' rather than 'caterpillar', and *pášni̱* 'pig' would be expected to mean 'bathed (one)' — which it does in appropriate contexts, where the referent is clearly not a pig. Such expressions are not compositional in any sense, and none of the elements are chosen freely: the entire morphological complex is chosen wholesale as a phraseologized chunk based on the meaning of the expression as a whole.

   As a bit of an aside, it should be pointed out that the failure to recognize phraseologization inside such morphological formations as these often leads, especially in Americanist studies (Holden 2009), to confusion in the interlinear

glosses of word forms. There is a strong tendency among researchers in "exotic" and under-documented languages to treat each recognizable portion of a word form as a sign and gloss it according to its inherent meaning (as determined by its appearance in other, unrestricted, contexts), even if these signs taken together in fact constitute an idiosyncratic phraseologized complex. Consider, for example, the following word form from Mohawk:

(25)   Mohawk
t–a–yoti–'nikù:r–v'ne
change–PAST–IT/THEM–mind–fall
'it shocked them'
(Mithun 1984 cited in Holden 2009: 161)

A literal gloss of the constituent parts of this word form, given in the 2nd line of (25), does not make clear their combined meaning, whereas an alternative approach, which recognizes the one-off, phraseologized nature of the combination of the elements *t-*, *'nikù:r*, and *v'ne* as a morphological idiom meaning 'shock someone' gives us a better picture of the nature of the expression, which could (and should) be glossed alternatively as:

(26)   *t–a–yoti–'nikù:rv'ne*
shock–PAST–IT/THEM–shock
'it shocked them'

Failure to take such an approach often creates unnecessarily complicated descriptions, which obscures the real state of affairs. It may be useful for some purposes to indicate the literal meaning of the components of idioms, including morphological idioms, but one has to strictly distinguish between actual conventional meanings and literal glosses. Specifying all inherent meanings while ignoring phraseologization amounts to giving etymological breakdowns instead of synchronic parses, and would be the equivalent of glossing words such as *indication* as [IN–DIC]–ATION or *restaurant* as [RE–STAUR]–ANT.[27]

Strong affixal derivational idioms are combinations of two derivational morphemes that, taken together, express a meaning that includes neither of their inherent signifieds. We have not found any of these in UNT, but examples can be taken from the Salishan language Lushootseed, which has a set of circumfixes formed from the prefix *dxʷ-* 'contained' and various members of the set of what are referred to in Salishan studies as "lexical suffixes" (essentially, derivational suffixes with "substantive" meanings). Three of these are given in (27):

(27)   *dxʷ–X–igʷəd* 'frame   <   *dxʷ-* 'contained'   *-igʷəd* 'interior of
          of mind'   body'

> *dxʷ–X–utsid* 'language'   <   *dxʷ-* 'contained'   *-utsid* 'mouth,
> opening'[28]
> *dxʷ–X–us* 'face'   <   *dxʷ-* 'contained'   *-us* 'upper body'
> (Hess 1998: 18–19):

These derivational affix-complexes are applied quite productively to different roots (generally, verbs) in the derivation of forms such as those given in (28):

(28)   *dxʷx̌ʷal'igʷəd* 'give up, resign oneself to defeat'
        <   X = *x̌ʷal'* 'lack control, be helpless'
      dxʷləli?utsid 'be a different language'
        <   X = *ləli?* 'be different'
      dxʷc'agʷusəb 'wash one's face'
        <   X = *c'agʷ* 'be washed'   *-əb* 'middle voice'[29]

Thus, while each circumfix has a clear meaning associated with it, this meaning does not seem to include the meanings of either of its components, although there are obvious figurative links between those meanings and the meanings of the circumfixes. Indeed, it might be possible to argue that the circumfixes are, in fact, weak idioms if we accept that, for instance, in Lushootseed the meaning 'language' can be semantically decomposed into 'COMMUNCATION SYSTEM contained in the mouth' or 'frame of mind' into 'EMOTIONAL STATE contained inside the body'. This is, of course, the same problem mentioned earlier with respect to phrasal idioms (see Note 9): defining their type correctly depends on establishing careful lexicographic definitions of noncompositional expressions, be they phrasal or morphological.

3.2.4.   *Noncompositional inflectional phrasemes: Inflectional idioms.*   An inflectional morphological idiom is a noncompositional restricted combination of inflectional affixal morphemes (that is, an affix-complex) that expresses an inflectional meaning; the individual affixes inside this combination have their own nonrestricted uses in other parts of the inflectional system. Inflectional idioms, however, differ from other types of idiom in two important respects. In the first place, inflectional idioms express meanings which are inflectional values or combinations of inflectional values of the stems to which they attach. Meanings of this type are different from lexical or derivational meanings in that the component meanings of an inflectional idiom always bear individually on the stem rather than on each other,[30] and as such the meanings of inflectional idioms can not be said to have a semantic pivot, giving us a fourth, uniquely morphological, type of idiom to consider — the *pivotless idiom*. With other types of idiom, the semantic pivot is expressed by the idiom as a whole, rather than by any of the idiom's components (the pivot is an "added" meaning); with pivotless idioms there is simply no pivot to be expressed. Another

difference between pivotless and other types of idiom is that the former seem not to be divisible into the same types of weak, semi-, and strong idioms that characterized the latter. This seems to fall out from the nature of inflectional meanings, an issue that we take up again in the last section of this paper.

A good example of an inflectional idiom in UNT is that in (29c), repeated from (1) above, which illustrates the marking of the first-person plural exclusive subject:

(29) a. *ikłtatáː*
         ik–łtatáː–yaː
         $1SG_{SUB}$–sleep–INCOMP
         'I sleep'

      b. *łtataːyáːw*
         łtatáː–yaː–w
         sleep–INCOMP–$1PL_{SUB}$:INCL
         'we-including-you sleep'

      c. *ikłtataːyáːw*
         ik–łtatáː–yaː–w
         $1SG_{SUB}$–sleep–INCOMP–$1PL_{SUB}$:INCL
         'we-excluding-you sleep' [= ŁTATÁː$_{INCOMP, 1PL.SUB:EXCL}$]

The first two word forms — in (29a)–(29b) — are free morphological expressions and correspond to the straightforward morphological rules given in (30):

(30) $1_{SUB} \oplus SG_{SUB} \Leftrightarrow \{1SG_{SUB}\}$ [= (29a)]
      $1_{SUB} \oplus PL_{SUB} \oplus INCLUSIVE \Leftrightarrow \{1PL_{SUB}:INCL\}$ [= (29b)]

The first rule maps the value 'first-person' of the inflectional category person of subject (i.e., the inflectional value or grammeme 1_{SUB}) and the value 'singular' of the category number of subject (SG_{SUB}) onto a single cumulative morpheme that expresses their combined values. In the second case, the morpheme $\{1PL_{SUB}:INCL\}$ cumulatively expresses the values of three categories — person of subject, number of subject, and exclusivity.[31] However, for the third word form (29c), a less-straightforward rule is necessary:

(31) $1_{SUB} \oplus PL_{SUB} \oplus EXCLUSIVE \Leftrightarrow \{1SG_{SUB}\} \oplus \{1PL_{SUB}:INCL\} \mid person_{OBJ} \neq 2$[32]

Here, neither of the constituent morphemes corresponds exactly to any particular inflectional value or combination of values. The $\{1PL_{SUB}:INCL\}$ morpheme partially corresponds to 1_{SUB} and PL_{SUB} in this word form, but does not express, as it usually does, INCLUSIVE. $\{1SG_{SUB}\}$ partially corresponds to 1_{SUB} but does not have one component of its inherent signified, SG_{SUB}: this meaning is not part of the meaning of the affix-complex *ik-* $\oplus$ *-w* at all. The value EXCLUSIVE has no individual form of its own, but rather is expressed by the combination of morphemes $\{1SG_{SUB}\}$ and $\{1PL_{SUB}:INCL\}$. Thus, the word form in (29c) is non-

compositional, and none of the elements expressing the number or person of the subject are freely chosen or have their inherent signifieds, making this a clear example of an inflectional idiom.

Another inflectional idiom is the affix-complex *ik-* $\oplus$ *ka:-* $\oplus$ *-n* in the word form in (32c), which shows the inflection of UNT verbs with first-person subjects and second-person objects where one or both of the subject and object is plural:

(32)   a.   *ika:túkslį*
           ik–Ø–ka:–tų̣ks–lį
           1SG$_{SUB}$–3$_{OBJ}$–PL$_{OBJ}$–hit–COMP
           'I hit$_{COMP}$ them'

       b.   *iktúksnį*
           ik–Ø–tų̣ks–n–lį
           1SG$_{SUB}$–SG$_{OBJ}$–hit–2$_{OBJ}$–COMP
           'I hit$_{COMP}$ you$_{SG}$'

       c.   *ika:túksnį*
           ik–ka:–tų̣ks–n–lį
           1SG$_{SUB}$–PL$_{OBJ}$–hit–2$_{OBJ}$–COMP
           (i)   'I hit$_{COMP}$ you$_{PL}$'
           (ii)   'we$_{EXCL}$ hit$_{COMP}$ you$_{SG}$'
           (iii)   'we$_{EXCL}$ hit$_{COMP}$ you$_{PL}$'

The form in (32c) is three-way ambiguous. The gloss in (32c-i), 'I hit$_{COMP}$ you$_{PL}$', is the "expected" (i.e., nonphraseologized) meaning. The corresponding word form is compositional and can be constructed by the regular rules for the expression of the individual values of the object-person and object-number categories, given in (11) above and repeated here for convenience in (33) along with an additional rule for the expression of the first-person singular subject:

(33) $2_{OBJ} \Leftrightarrow \{2_{OBJ}\}$
       $SG_{OBJ} \Leftrightarrow \{SG_{OBJ}\}$
       $PL_{OBJ} \Leftrightarrow \{PL_{OBJ}\}$
       $1_{SUB} \oplus SG_{SUB} \Leftrightarrow \{1SG_{SUB}\}$

All of these rules are unrestricted and are generally applicable in other areas of the inflectional system. The glosses in (32c-ii) and (32c-iii), however, are "unexpected" and require additional specialized rules. The form with the gloss 'we$_{EXCL}$ hit$_{COMP}$ you$_{SG}$' (32c-ii) would be described by the rule for the expression of the second-person object in (33) plus the rule in (34):

(34) $1_{SUB} \oplus PL_{SUB} \oplus SG_{OBJ} \oplus$ EXCLUSIVE $\Leftrightarrow \{1SG_{SUB}\} \oplus \{PL_{OBJ}\} \mid$ person$_{OBJ} = 2$

According to this rule, when the object of the verb is second person, the combination of the inflectional values $1_{SUB} \oplus PL_{SUB} \oplus SG_{OBJ} \oplus$ EXCLUSIVE is realized

by the combination of morphemes $\{1\text{SG}_{\text{SUB}}\} \oplus \{\text{PL}_{\text{OBJ}}\}$. In much the same way, the form with the gloss 'we$_{\text{EXCL}}$ hit$_{\text{COMP}}$ you$_{\text{PL}}$' (32c-iii) would be described by the rule for the expressions of the second-person object and the plural object in (33), as well as the rule in (35):

(35) $1_{\text{SUB}} \oplus \text{PL}_{\text{SUB}} \oplus \text{EXCLUSIVE} \Leftrightarrow \{1\text{SG}_{\text{SUB}}\} \mid \text{person}_{\text{OBJ}} = 2$

Note that the left side of this rule (that is, the set of inflectional values expressed by this rule) is the same as that given in (31) for the expression of the first-person plural exclusive subject, but the right side (the morpheme that corresponds to these inflectional values) is different, as are the conditions on its application: whereas the rule in (31) applies when the person of the object is not 2, the rule in (35) applies when the person of the object is 2. The morpheme that is selected by this rule, $\{1\text{SG}_{\text{SUB}}\}$, is not a compositional expression of the inflectional values it corresponds to, nor does it express a semantic pivot. The same applies in (34), which, like (35), describes an inflectional idiom.

A similar case is found in verb forms where the subject is second-person, the object first-person, and one or both of them is plural, as shown in (36d):

(36)   a.   *kintúksli̱*
                kin–Ø–Ø–tu̱ks–li̱
                1_{OBJ}–SG$_{\text{OBJ}}$–3SG$_{\text{SUB}}$–hit–COMP
                's/he hit$_{\text{COMP}}$ me'

          b.   *túkswi̱*
                Ø–Ø–tu̱ks–w–li̱
                3_{OBJ}–SG$_{\text{OBJ}}$–hit–1PL$_{\text{SUB}}$:INCL–COMP
                'we$_{\text{INCL}}$ hit$_{\text{COMP}}$ him/her'

          c.   *la:túkswi̱*
                la:–tu̱ks–w–li̱
                RCP–hit–1PL$_{\text{SUB}}$:INCL–COMP
                'we$_{\text{INCL}}$ hit$_{\text{COMP}}$ each other'

          d.   *kila:túkswi̱*
                kin–la:–tu̱ks–w–li̱
                1_{OBJ}–RCP–hit–1PL$_{\text{SUB}}$:INCL–COMP
                (i)   'you$_{\text{SG}}$ hit$_{\text{COMP}}$ us'
                (ii)   'you$_{\text{PL}}$ hit$_{\text{COMP}}$ us'
                (iii) 'you$_{\text{PL}}$ hit$_{\text{COMP}}$ me'

The word form in (36d) is completely unexpected, containing the affixes that normally express a first-person plural subject and a first-person singular object, as well as the reciprocal prefix, *la:-*. This type of form is a clear case of what Baerman et al. (2007) refer to as "deponency," the presence of the first-person plural subject markers (and the reciprocal marker) in a sense "contradicting"

the actual meaning of the form. In addition to the rule for the expression of the first-person object, the description of this word form requires another set of special rules, given in (37), which map specific configurations of inflectional values onto the combination of morphemes $\{RCP\} \oplus \{1PL_{SUB}\}$:

(37)   a. $2_{SUB} \oplus SG_{SUB} \oplus PL_{OBJ} \Leftrightarrow \{RCP\} \oplus \{1PL_{SUB}:INCL\}$ | person$_{obj}$ = 1
           [= (36d-i)]
       b. $2_{SUB} \oplus PL_{SUB} \oplus PL_{OBJ} \Leftrightarrow \{RCP\} \oplus \{1PL_{SUB}:INCL\}$ | person$_{obj}$ = 1
           [= (36d-ii)]
       c. $2_{SUB} \oplus PL_{SUB} \oplus SG_{OBJ} \Leftrightarrow \{RCP\} \oplus \{1PL_{SUB}:INCL\}$ | person$_{obj}$ = 1
           [= (36d-iii)]

For each of these rules, the condition on its application is the presence of a first-person object. The various configurations of inflectional values on the left side of the rules find their expressions in the combination of the first-person object marker with the $\{RCP\}$ and $\{1PL_{SUB}:INCL\}$ morphemes, neither of which correspond directly to any part of their meanings. Thus, all of these expressions are noncompositional inflectional affix-complexes — that is, inflectional idioms.

   Additional examples of inflectional idioms can be found in the UNT inflections for grammatical moods, one of which, the irrealis, has no dedicated affixes of its own, but instead "borrows" from the affixes used for tense, aspect, and other moods to create noncompositional affix-complexes for its expression. The irrealis mood indicates that an event that might have been realized at the time of reference has not been and never will be realized; this mood is incompatible with aspect, but is (except in the past irrealis) expressed cumulatively with tense.

   The present tense of the irrealis mood is expressed by a combination of the completive suffix and the past-tense prefix, as shown in (38c):

(38)   a.   *pasá:ł*
           Ø–Ø–pasá:–l$i̠$
           PRES–3SG$_{SUB}$–burn–COMP
           'it burns$_{COMP}$ up'
       b.   *išpasá:*
           iš–Ø–pasá:–ya:
           PAST–3SG$_{SUB}$–burn–INCOMP
           'it burned/was burning'
       c.   *išpasá:ł*
           iš–Ø–pasá:–l$i̠$
           PAST–3SG$_{SUB}$–burn–COMP
           'had it burned/it could have burned' [= PASÁ:$_{IRREAL, PRES}$]

The rule that describes this inflectional pattern is given in (39); it shows clearly that there is no correspondence at all between the inflectional values being expressed and the morphemes being used to express them:

(39)   IRREALIS ⊕ PRESENT ⟺ {PAST} ⊕ {COMP}

The expression of the present irrealis is thus both restricted in that it is selected as a whole, based on the desired tense-mood combination (present plus irrealis), and noncompositional.

Similarly, the future tense of the irrealis mood is expressed by a combination of the completive suffix, the potential prefix *ti-*, and the optative prefix *ka-*:

(40)   a.   *pasá:ł*
             Ø–Ø–pasá:–lị
             PRES–3SG_SUB–burn–COMP
             'it burns_COMP up'
       b.   *tipasá:*
             Ø–ti–Ø–pasá:–ya:
             PRES–POT–3SG_SUB–burn–INCOMP
             'it could/might burn'
       c.   *kapasá:*
             ka–Ø–pasá:–ya:
             OPT–3SG_SUB–burn–INCOMP
             'let it burn!'
       d.   *katipasá:ł*
             ka–Ø–ti–pasá:–lị
             OPT–3SG_SUB–POT–burn–COMP
             'it could catch on fire/burn (but won't now)' [= PASÁ:_IRREAL, FUT]

In the expression of the future irrealis, we see the borrowing of affixes from different moods to form part of an inflectional idiom. The inherent meanings of the potential *ti-* (used for events which are unrealized at the reference time, but are potentially realizable) and the optative *ka-* (used for desired states of affairs, including imperatives and hortatives) are not part of the meaning of the affix-complex as a whole, FUTURE IRREALIS, nor is the meaning of the completive suffix. As before, the rule describing this inflection, shown in (41), reveals a complete mismatch between inflectional values and their morphemic expressions:

(41)   IRREALIS ⊕ FUTURE ⟺ {OPT} ⊕ {POT} ⊕ {COMP}

The affix-complex described by this rule is also noncompositional and contains no semantic pivot, making it another example of an inflectional idiom.

In the third member of the mood paradigm, the past irrealis, the irrealis is expressed by the combination of the completive suffix and the potential prefix,

but unlike the other irrealis inflections it contains a separate marker of past tense actually signifying PAST:

(42)   a.   *tipasá:ł*
             Ø–Ø–ti–pasá:–li̠
             PRES–3SG_{SUB}–POT–burn–COMP
             'it can burn_{COMP} up'
       b.   *ištipasá:*
             iš–Ø–ti–pasá:–ya:
             PAST–3SG_{SUB}–POT–burn–INCOMP
             'it could burn (and maybe it did later)'
       c.   *ištipasá:ł*
             iš–Ø–ti–pasá:–li̠
             PAST–3SG_{SUB}–POT–burn–COMP
             'it could have burned (but didn't)' [= PASA:_{IRREAL, PAST}]

The expression of the past tense of the irrealis mood contains the morpheme {PAST}, which corresponds to the inflectional value PAST in its meaning (and which would be covered by a general rule for the expression of the past tense, PAST ⇔ {PAST}). The other morphemes, {POT} and {COMP}, however, do not correspond to a particular value of any inflectional category, and form an affix-complex which is selected as a whole for the expression of the irrealis mood itself if and only if the verb is also inflected for the past tense, as described by the following rule:

(43)   IRREALIS ⇔ {POT} ⊕ {COMP} | tense = PAST

Because the inherent signifieds of neither morpheme correspond to IRREALIS, the rule in (43) describes a pivotless idiom whose components are not chosen separately based on their individual signifieds but instead are selected as a "chunk" based on the Speaker's desired meaning (both in terms of the desired mood and the tense of the expression).

    As noted above in the discussion of the progressive aspect marker (Subsection 3.2.1, Note 23), it is logically possible to treat expressions such as those in (38), (40), and (42) as morphemes in their own right (that is, as irrealis circumfixes); however, this seems like a poor option in terms of descriptive and formal economy: the allomorphic, ordering, and other formal properties of the morphemes that constitute the irrealis mood-markers are precisely the same in their free and in their restricted uses. Thus, to take only one example, the completive suffix in the present irrealis form shows the same alternations due to verbal inflection class that the completive suffix in the indicative (and other) moods shows:

(44)   a.   *ikmusú:ł*
             ik–Ø–Ø–musú:–li̠
             1SG_{SUB}–3_{OBJ}–SG_{OBJ}–kiss–COMP
             'I kissed him/her'

       b.   *šakmusú:ł*
             iš–ik–Ø–Ø–musú:–li̠
             PAST–1SG_{SUB}–3_{OBJ}–SG_{OBJ}–kiss–COMP
             'had I kissed him/her'

       c.   *iktṵ́ksli̠*
             ik–Ø–Ø–tṵks–li̠
             1SG_{SUB}–3_{OBJ}–SG_{OBJ}–hit–COMP
             'I hit it'

       d.   *šaktṵ́ksli̠*
             iš–ik–Ø–Ø–tṵks–li̠
             PAST–1SG_{SUB}–3_{OBJ}–SG_{OBJ}–hit–COMP
             'had I hit it'

       e.   *iklạʔtsíł*
             ik–Ø–Ø–lạʔtsín–li̠
             1SG_{SUB}–3_{OBJ}–SG_{OBJ}–see–COMP
             'I saw it'

       f.   *šaklạʔtsíł*
             iš–ik–Ø–Ø–lạʔtsín–li̠
             PAST–1SG_{SUB}–3_{OBJ}–SG_{OBJ}–see–COMP
             'had I seen it'

UNT verb stems fall into three general inflection classes, represented in (44) by
MUSÚ: 'kiss someone' (Class I), TṴKS- 'hit something' (Class II), and LAʔTSÍN
'see something' (Class III). In the first- and third-persons, Class I and Class III
verbs take the *-ł* allomorph of the completive suffix, whereas Class II verbs
take the full form of the suffix, *-li̠*. In the second-person singular, Class I and
III mark their subjects through leftward movement of stress and laryngealiza-
tion of the final vowel (which also cumulatively marks the completive aspect),
while Class II verbs take a suffix *-ti̠*:

(45)   a.   *músu̠*
             Ø–Ø–músu̠
             3_{OBJ}–SG_{OBJ}–kiss:2SG_{SUB}:COMP
             'you_{SG} kissed him/her'

       b.   *išmúsu̠*
             iš–Ø–Ø–músu̠
             PAST–3_{OBJ}–SG_{OBJ}–kiss:2SG_{SUB}:COMP
             'had you_{SG} kissed him/her'

c. *túkstį*
Ø–Ø–tṵks–t–lį
3_{OBJ}–SG_{OBJ}–hit–$2SG_{SUB}$–COMP
'you$_{SG}$ hit it'

d. *ištúkstį*
iš–Ø–Ø–tṵks–t–lį
PAST–3_{OBJ}–SG_{OBJ}–hit–$2SG_{SUB}$–COMP
'had you$_{SG}$ hit it'

e. *láʔtsį*
Ø–Ø–láʔtsį
3_{OBJ}–SG_{OBJ}–see:$2SG_{SUB}$:COMP
'you$_{SG}$ saw it'

f. *išláʔtsį*
iš–Ø–Ø–láʔtsį
PAST–3_{OBJ}–SG_{OBJ}–see:$2SG_{SUB}$:COMP
'had you$_{SG}$ seen it'

In the same vein, first-person plural completive forms of Class II verbs take the suffix *-wį*, and transitive Class II verbs with second-person objects in the completive aspect take the suffix *-nį*. Class I and III verbs, on the other hand, truncate the completive marker completely in both these contexts (i.e., *-w* + *-lį* ⇒ *-w*; *-n* + *-lį* ⇒ *-n*). Other behaviors shared by the free and restricted uses of the completive suffix include its interaction with the indefinite voice marker and the repetitive suffix (see Note 23), its effects on the realization of the debitative marker, and its morphophonemic interaction with the stative plural marker.

If, in a formal morphological model, the *-lį* suffixes in the three irrealis forms are treated as different morphemes, this would require that each of the rules describing the behavior of the completive affixes be repeated four times (once for the unrestricted use and once for each of the three irrealis forms). The economy of treating this suffix as part of a morphological idiom is thus considerable. In the formal model of UNT nonderivational morphology proposed by Beck et al. (n.d.), for instance, the completive morpheme appears in the input to, or in the conditions on, 12 separate morphological and morphophonological rules. Treating the *-lį* suffix as part of three different irrealis circumfixes would increase the number of rules needed to 48 (12 for *-lį* itself plus 12 × 3 for each occurrence of *-lį* as part of one of the irrealis circumfixes). Similar considerations apply to the other affixes involved in irrealis marking (which have their own, albeit less complicated, combinatorial and allomorphic properties),[33] as well as to the affixes involved in the other morphological idioms and collocations discussed above. All told, the total reduction in the number of rules needed to describe UNT verbal morphology afforded by treating these

expressions as morphological phrasemes composed of morphemes that are freely chosen in other contexts runs into the hundreds (see Beck 2007 for further discussion).

4.   Conclusions

This paper has attempted to show that phraseologization, well-known at the level of phrases, where it produces restricted multi-word expressions, also operates at the level of words, giving rise to single-word or smaller expressions made up of phraseologized combinations of radical and/or affixal morphemes (and, potentially, nonconcatenative morphological operations as well). We began by sketching a theory of phraseology, defining the phraseologization of complex linguistic signs, identifying their taxonomic properties (paradigmatic and syntagmatic freedom), and establishing the major types of phraseologized expression these two properties can be used to define. Within each of these types of phraseme, compositional and noncompositional, we further distinguished specific subtypes, as summarized in Table 2. In the subsequent sections of the paper, we identified morphological expressions — compound, derivational and inflectional — that correspond to most of these subtypes, which hitherto had only been recognized on the phrasal level. These findings are summarized in Table 6. The strong parallels between phraseologized expressions on the syntactic (phrasal) and morphological (morphemic) level lead us to conclude that morphological phrasemes are, in general, amenable to the same classifications as phrasal phrasemes — but with some exceptions: lexical-syntactic and morphological signs are after all of somewhat different natures. In the first place, the number of words in a language is finite, while the number of phrases is potentially infinite. This entails a much higher variability among phrases as compared to the variability of the set of words. As a consequence, two of the phraseme-types identified in Table 2 at the phrasal level, pragmatemes and clichés, have yet to be found on the morphological level. Both cases require a certain amount of synonymy between expressions so that pragmatic or conventional factors can come into play and restrict the selection of one semantically (near-)equivalent form over another. Given the relative scarcity of morphological resources (numbered at the most in the hundreds) as compared to lexical resources (numbered in the hundreds of thousands), such synonymy is likely to be uncommon, and pragmatic and conventional restrictions on the selection of synonymous expressions even more so. At this point it might be premature to rule out the possibility of morphological pragmatemes and morphological clichés completely, given that this line of research has only just begun, but there do seem to be good reasons not to expect to see many examples of these two types of morphological phraseme.

Table 6.   *Types of phrasemes*

		phrasal	compound	derivational		inflectional	
				full	affixal	full	affixal
compositional	collocation	✓	✓	✓	✓	✓	✓
	pragmateme	✓	?	?	?	?	?
	cliché	✓	?	?	?	?	?
noncompositional	weak idiom	✓	✓	✓	✓	✗	
	semi-idiom	✓	✓	✓	✓	✗	pivotless
	strong idiom	✓	✓	✓	✓	✗	idiom

✓ = attested; ✗ = not possible; ? = as yet unattested

A further observation from Table 6 is that, in general, the parallels between phrasemes on the morphological and lexical level are stronger for compound and derivational phrasemes than for inflectional phrasemes. This is also not surprising in that compounding and derivation (both types of word formation) add new lexical units (LUs) to the lexicon in precisely the same way that the coining of new phrasal idioms adds new (multi-word) LUs. Stored multiword LUs (i.e., idioms) are associated with meanings that are not the regular sum of their component forms; the same is true of idiomatic compounds and derivatives, which also constitute LUs. Compound and derivational collocations are likewise parallel to lexical collocations, which are described in the lexical entries of their bases. For compositional inflectional phrasemes, the situation is quite similar. With full (i.e., stem + affix) inflectional collocations, each inflected word form is associated with an LU, and that association can be conventionalized such that for any particular LU (say, a Burushaski noun) requiring some inflectional category, a particular morphological marker of that category (a particular plural suffixal morpheme) is specified for that LU. Affixal compositional inflectional phrasemes (inflectional collocations or parasitic formations) also consist of a conventionalized association between one element (in this case an inflectional affix) and a collocate which must be learned and stored by a Speaker, although they represent a somewhat different type of phraseme from other collocations in that their base does not express their semantic pivot — indeed, they have no semantic pivot in their meaning at all, being pivotless collocations.

When it comes to noncompositional inflectional phrasemes — that is, inflectional idioms, which are also pivotless — the parallelism with lexical phrases breaks down in another respect: inflectional idioms cannot be subdivided into weak, semi-, and strong subtypes. We believe that this stems from the special nature of inflectional meanings. Unlike the meanings of derivational affixes, the meanings of inflectional affixes are highly abstract (occasionally to the

point of being purely formal), form paradigmatic sets, and are amenable to descriptions by rules couched in terms of the combination of values of inflectional categories that may not correspond semantically to the real-world situation being expressed (for example, English *pluralia tantum* or the historical present tense).[34] Likewise, the meanings of inflectional idioms must be composed of combinations of such abstract meanings, and thus expressible in terms of combinations of values of inflectional categories. Because the definitions of two of the subtypes of idiom — the weak idiom and the semi-idiom — specify that the meaning of the whole contain the meaning of all or some of its parts, an inflectional idiom of either type would express the inflectional values that some or all of the constituent morphemes inherently express. This makes that morpheme "extractable" from the idiom and allows us to treat it as a separate, regular marker of the relevant values. This kind of reduction is simply not an option for derivational idioms, where derivational meanings are not discrete, but bear on one another to form semantically complex expressions whose meanings are not always exclusively decomposable into paradigmatic values in the same way that inflections are. Stated in another way, in terms of rules, if any morpheme on the right side of a rule describing an inflectional idiom can be shown to be solely responsible for the expression of a particular inflectional value on the left side of that rule, then it should be possible to write a separate more general rule for this correspondence, reducing the inflectional idiom to the morphemes that do not show such correspondences. This procedure reduces all candidates for inflectional semi-idioms to weak idioms. Weak inflectional idioms, in turn, can be eliminated by assigning each of their component morphemes that morpheme's inherent meaning and positing a zero exponent for any of the remaining inflectional values. Thus, the only remaining possibility is the strong inflectional idiom, in which the inherent meaning of none of the components is part of the meaning of the whole. In a technical sense, this is a correct description of the inflectional idiom; however, inflectional idioms (unlike other strong idioms) do not have a semantic pivot, and we have chosen to refer to them as "pivotless idioms" rather than as "strong inflectional idioms" in recognition of that fact.

Despite the differences between morphological and phrasal phrasemes, it seems clear from the data presented in this paper that the general principles of phraseology — that is, the principles governing the restricted combination of linguistic signs and/or sets of signs — apply equally to phrases and to morphologically-complex words. This seems to us to be strong evidence for the utility of treating the constituent parts of complex words, morphs, as linguistic signs: simply put, if it walks like a sign and quacks like a sign, it must be a sign. Describing the behavior and meanings of phraseologized combinations of parts of words is difficult without recognizing those parts as separable and analyzable elements to begin with. The fact that the combinatoric proper-

ties of individual parts of these phraseologized strings of affixal morphemes are identical for both their free and their restricted forms is a good indication that the two occurrences of the morph are the "same" in the sense of sharing a representation (or being representationally linked) in the minds of speakers. Certainly, in terms of the economy of formal representation (and, perhaps, in terms of learnability), the best analytical choice is to treat the occurrence of the affix found in the phraseme as a restricted use of the affix that is found freely elsewhere.

It is, of course, not the claim of this paper that other, word-based approaches to morphology are unable to handle the particular descriptive facts that we have outlined in this paper — after all, as we noted in our introduction, these theories were designed precisely to deal with such phenomena. Indeed, the concept of the morpheme as we have presented it here bears a strong resemblance (if not a functional equivalence) to a paradigm cell in a Word and Paradigm model (Matthews 1991), and the rules we present mapping between grammemes and morphemes in order to allow for "mismatches" between values of inflectional categories and the morphemes that express them achieve results similar to the "rules of referral" (Zwicky 1985) employed in such models. However, we feel that an essential difference between the two approaches resides in the fact that word-based theories treat what we are calling morphs as essentially empty phonological elements that have in themselves no corresponding meaning, whereas we continue to treat sub-lexical units, morphs, as linguistic signs. By not doing so, word-based approaches would seem to have no way to account for why meaningless substrings of words should show the same types of phraseological behavior that meaningful elements like words do. In effect, we suggest that rather than completely abolishing the morph and the morpheme, revising them along the lines suggested here allows us to resolve morphological problems of the type outlined in this paper without giving up the utility of the morph, morpheme, and linguistic sign.

While the observation that lies at the heart of this paper — that morphs are signs and, like words, participate in phraseologization — may be out of tune with some of the more recent trends in morphology, it seems very much in step with many contemporary usage-based approaches to syntax (in particular cognitive and construction grammars — e.g., Langacker 1987, 1991; Fillmore et al. 1988; Goldberg 1995), which have placed a great deal of emphasis on what we term phrasal phrasemes (a.k.a. "constructions") and their role in syntax. Extending the notion of "construction" to the realm of morphology seems like a natural step, and is one that is implicit in much of the work being carried on the meanings of words and affixes in these frameworks. However, unlike some of the more radical versions of these theories (e.g., Croft 2001), the approach taken here does not require us to abandon the notion of the sub-units or constituents of constructions as meaningful elements in and of themselves.

Instead, it allows us to treat them as elements with inherent meanings (linguistic signs) that can take on restricted and contingent meanings, or combine to form complex signs with entirely different meanings under specified conditions. This is not a radical idea, but it is perhaps one which has not been pursued to its ultimate conclusions. We hope that the utility of doing so, particularly in the realm of morphology, has become clear throughout the discussion in this paper. Certainly from the point of view of detailed and accurate linguistic description, as well as from the perspective of formal modeling and text synthesis, the approach to morphological (and other types of) phraseologization outlined here has some obvious advantages, and seems to reflect important properties of natural language.

Appendix. Upper Necaxa verbal paradigms

Presented below are some sample person, aspect, and mood forms for the Upper Necaxa verbs used in this paper to allow the interested reader to see the complete paradigms to which they belong. Regular dynamic verbs in Upper Necaxa belong to one of three classes — Class I including all vowel-final stems, Class II consisting of most consonant-final stems, and Class III being an idiosyncratic set of n-final stems.

Table A1.   *Present-tense person and aspect forms of* ɬatá: *'sleep' (Class I)*

	INCOMPLETIVE	COMPLETIVE	PERFECT	PROGRESSIVE
1SG	*ikɬatá:*	*ikɬatá:ɬ*	*ikɬata:ní̠:*	*ikɬata:má:ɬ*
2SG	*ɬatá:ya̠*	*ɬátạ:*	*ɬata:ní̠ta̠*	*ɬata:pá̠:*
3SG	*ɬatá:*	*ɬatá:ɬ*	*ɬata:ní̠:*	*ɬata:má:ɬ*
1PL.EXCL	*ikɬata:yá:w*	*ikɬatá:w*	*ikɬata:ní̠:táw*	*ikɬata:ma̠:náw*
1PL.INCL	*ɬata:yá:w*	*ɬatá:w*	*ɬata:ní̠:táw*	*ɬata:ma̠:náw*
2PL	*ɬata:ya:tít*	*ɬata:tít*	*ɬata:ní̠:tantít*	*ɬata:pa̠:nantít*
3PL	*taɬtatá:*	*taɬtatá:ɬ*	*taɬtata:ní̠:*	*taɬtata:ma̠:náɬ*

Table A2.   *Present incompletive person forms of* musú: *'kiss someone' (Class I)*

OBJECT / SUBJECT	1SG	2SG	3SG
1SG	—	*ikmusu:yá:n*	*ikmusú:*
2SG	*kimusú:ya̧*	—	*musú:ya̧*
3SG	*kimusú:*	*musu:yá:n*	*musú:*
1PL.EXCL	—	*ika:musu:yá:n*	*ikmusu:yá:w*
1PL.INCL	—	—	*musu:yá:w*
2PL	*kila:musu:yá:uw*	—	*musu:ya:tít*
3PL	*kintamusú:*	*tamusu:yá:n*	*tamusú:*

	1PL	2PL	3PL
1SG	—	*ika:musu:yá:n*	*ika:musú:*
2SG	*kila:musu:yá:uw*	—	*ka:musú:ya̧*
3SG	*kinka:musu:yá:n*	*ka:musu:yá:n*	*ka:musú:*
1PL.EXCL	—	*ika:musu:yá:n*	*ika:musu:yá:w*
1PL.INCL	—	—	*ka:musu:yá:w*
2PL	*kila:musu:yá:uw*	—	*ka:musu:ya:tít*
3PL	*kinka:tamusu:yá:n*	*ka:tamusu:yá:n*	*tamusú:/ka:musú:* *

*Plurality of subject and object cannot be marked simultaneously for 3rd persons.

Table A3.   *Present completive person forms of* musú: *'kiss someone' (Class I)*

OBJECT / SUBJECT	1SG	2SG	3SG
1SG	—	*ikmusú:n*	*ikmusú:ł*
2SG	*kimúsu̧:*	—	*músu̧:*
3SG	*kimusú:ł*	*musú:n*	*musú:ł*
1PL.EXCL	—	*ika:musú:n*	*ikmusú:w*
1PL.INCL	—	—	*musú:w*
2PL	*kila:musú:w*	—	*musu:tít*
3PL	*kintamusú:ł*	*tamusú:n*	*tamusú:ł*

	1PL	2PL	3PL
1SG	—	*ika:musú:n*	*ika:musú:ł*
2SG	*kila:musú:w*	—	*ka:músu̧:*
3SG	*kinka:musú:n*	*ka:musú:n*	*ka:musú:*
1PL.EXCL	—	*ika:musú:n*	*ika:musú:w*
1PL.INCL	—	—	*ka:musú:w*
2PL	*kila:musú:w*	—	*ka:musu:tít*
3PL	*kinka:tamusú:n*	*ka:tamusú:n*	*tamusú:ł/ka:musú:ł* *

*Plurality of subject and object cannot be marked simultaneously for 3rd persons.

Table A4.   *Present completive person forms of* tu̱ks- *'hit someone' (Class II)*

OBJECT / SUBJECT	1SG	2SG	3SG
1SG	—	*iktu̱ksni̱*	*iktu̱ksli̱*
2SG	*kintu̱ksti̱*	—	*tu̱ksti̱*
3SG	*kintu̱ksli̱*	*tu̱ksni̱*	*tu̱ksli̱*
1PL.EXCL	—	*ika:tu̱ksni̱*	*iktu̱kswi̱*
1PL.INCL	—	—	*tu̱kswi̱*
2PL	*kila:tu̱kswi̱*	—	*tu̱kstít*
3PL	*kintatu̱ksli̱*	*tatu̱ksni̱*	*tatu̱ksli̱*
	1PL	**2PL**	**3PL**
1SG	—	*ika:tu̱ksni̱*	*ika:tu̱ksli̱*
2SG	*kila:tu̱kswi̱*	—	*ka:tu̱ksti̱*
3SG	*kinka:tu̱ksni̱*	*ka:tu̱ksni̱*	*ka:tu̱ksli̱*
1PL.EXCL	—	*ika:tu̱ksni̱*	*ika:tu̱kswi̱*
1PL.INCL	—	—	*ka:tu̱kswi̱*
2PL	*kila:tu̱kswi̱*	—	*ka:tu̱kstít*
3PL	*kinka:tatu̱ksni̱*	*ka:tatu̱ksni̱*	*tatu̱ksli̱|ka:tu̱ksli̱* *

* Plurality of subject and object cannot be marked simultaneously for 3rd persons.

Table A5.   *Present completive person forms of* la̱ʔtsín *'see someone' (Class III)*

OBJECT / SUBJECT	1SG	2SG	3SG
1SG	—	*ikla̱ʔtsín*	*ikla̱ʔtsíɬ*
2SG	*kilá̱ʔtsi̱*	—	*lá̱ʔtsi̱*
3SG	*kila̱ʔtsíɬ*	*la̱ʔtsín*	*la̱ʔtsíɬ*
1PL.EXCL	—	*ika:la̱ʔtsín*	*ikla̱ʔtsíw*
1PL.INCL	—	—	*la̱ʔtsíw*
2PL	*kila:la̱ʔtsíw*	—	*la̱ʔtsintít*
3PL	*kintala̱ʔtsíɬ*	*tala̱ʔtsín*	*tala̱ʔtsíɬ*
	1PL	**2PL**	**3PL**
1SG	—	*ika:la̱ʔtsín*	*ika:la̱ʔtsíɬ*
2SG	*kila:la̱ʔtsiw*	—	*ka:lá̱ʔtsi̱*
3SG	*kinka:la̱ʔtsín*	*ka:la̱ʔtsín*	*ka:la̱ʔtsíɬ*
1PL.EXCL	—	*ika:la̱ʔtsín*	*ika:la̱ʔtsíw*
1PL.INCL	—	—	*ka:la̱ʔtsíw*
2PL	*kila:la̱ʔtsíw*	—	*ka:la̱ʔtsintít*
3PL	*kinka:tala̱ʔtsín*	*ka:tala̱ʔtsín*	*tala̱ʔtsín|ka:la̱ʔtsín* *

* Plurality of subject and object cannot be marked simultaneously for 3rd persons.

Table A6.   *3 SG subject TAM forms of* pasá: *'burn' (Class I)*

	IMPERFECTIVE	PERFECTIVE	PROGRESSIVE	PERFECT
INDICATIVE				
FUTURE	*napasá:*			
PRESENT	*pasá:*	*pasá:ł*	*pasa:má:ł*	*pasa:nį:*
PAST	*išpasá:*		*išpasa:má:ł*	*išpasa:nį:*
OPTATIVE				
FUTURE				
PRESENT	*kapasá:*	*kapasá:ł*	*kapasa:má:ł*	*kapasa:nį:*
PAST				
POTENTIAL				
FUTURE	*natipasá:*			
PRESENT	*tipasá:*	*tipasá:ł*	*tipasa:má:ł*	*tipasa:nį:*
PAST	*ištipasá:*		*ištipasa:má:ł*	*ištipasa:nį:*
IRREALIS				
FUTURE	*kapasá:ł*			
PRESENT	*išpasá:ł*			
PAST	*ištipasá:ł*			

Notes

1.   Upper Necaxa Totonac is a member of the Totonac-Tepehua family, spoken by some 3,400 people in the Sierra Norte of Puebla State, Mexico. Uncited data from this language are taken from Beck's fieldnotes. The abbreviations used in interlinear glosses are as follows: 1, 2, 3 = first-, second-, third-person; COMP = completive aspect; DAT = dative case; EXCL = exclusive; ERG = ergative case; GEN = genitive case; IMPF = imperfective aspect; INCL = inclusive; INCOMP = incompletive aspect; INDF = indefinite voice; OBJ = object; OPT = optative mood; PL = plural; POT = potential mood; PROG = progressive aspect; RCP = reciprocal; RPT = repetitive; SG = singular; SUB = subject; TOT = totalitative. The first analyzed line of a gloss presents the basic allomorphs of the corresponding morphemes, unaffected by morphophonemic processes. Here and in the remainder of the paper, we omit zero affixes, such as the markers of tense and mood in (1), in order not to clutter the exposition unless these affixes

are directly implicated in the particular point under discussion. Full inflectional paradigms for some Upper Necaxa verbs are given in Appendix I.

2. The incompletive suffix *-ya:* is truncated in final position here by a morphophonological process.

3. S. Kahane (p. c.) points out that there is a third approach that might apply to this data: to make the example in (1c) compositional by adopting a different, nonstandard approach to the inflectional categories of person and number — namely, by replacing 'first-person singular' with something like 'involving me but not you', 'second-person singular' with 'involving you, but not me', etc. However, in spite of considerable effort, we have so far failed to find a single alternative analysis that resolves the difficulties in (1c) and, at the same time, addresses all similar problems found in other areas in the system, discussed in the sections below (for example, in the marking of object-agreement, which does not make the inclusive-exclusive distinction but presents other complications of its own). Even were our efforts successful, we find the kind of formal fixes that would be needed to make forms like (1c) compositional unnatural and implausible as the putative basis for natural language systems — or, at any rate, less plausible than the appeal being made in this paper to the principles of phraseologization, which are taken for granted in other areas of the grammar. A full discussion of these issues and the extensive literature on the semantics of "exotic-looking" person-number systems (e.g., Hockett 1966; Harley and Ritter 2002; Cysouw 2003) is, of course, a topic for a different paper.

4. The "Speaker" is the author of a particular speech act, as opposed to a "speaker" of a language.

5. The bipartite division of semantic phrasemes into collocations and idioms goes back to the classic work of Bally (1950 [1932]: 66–87); on the subdivision of idioms, see Mel'čuk (1982: 118–119).

6. The raised corner brackets around ˹KICK THE BUCKET˺ are used to indicate that this is an idiom rather than a syntagmatically free compositional expression.

7. The base of the collocation is printed in SMALL CAPS, while the collocates are given in *italics*.

8. Another way to think of it is in terms of paraphrasing. Inside the meaning of a single-word expression (= a lexeme) L, the semantic pivot coincides with the minimal paraphrase of L such that replacing L with this paraphrase, you lose a part of the information carried by L, but you do not deform it. If GLASS is 'container designed for people to drink from . . .', then we can say CONTAINER instead of GLASS, and this will not be a lie, only less precise. The same applies to BLACK COFFEE: if you ask only for "coffee," you will get the right beverage, only it may not be prepared according to your taste.

9. Correctly distinguishing these idiom types naturally hinges on determining their correct lexicographical definitions, which is no mean feat. This problem is treated in Mel'čuk (2006b), but is too profound for us to pursue here. The exposition that follows will simply present what we believe to be the correct definitions of idioms without fully justifying them. Individual English examples are simply intended as heuristic devices for illustrating the taxonomic principles involved, and nothing crucial hinges on whether or not a particular example might have been misclassified.

10. Of course, the formal requirements of morpheme-hood must be defined more rigorously than is done here. For a detailed discussion, see Mel'čuk (1982: 88–91, 1993–2000: vol. 4, 201–228, 2006a).

11. We use the term "category" here to refer to a set of paradigmatically related values of some type of grammatical meaning such as tense, aspect, or person. In some works, an inflectional category is referred to as a "feature." However, it is our feeling that this term is linked too tightly to a type of syntactically-driven view of morphology that we do not share.

12. In the remainder of this paper we will be following the formal conventions of Meaning-Text Theory (Žolkovskij and Mel'čuk 1965, 1967; Mel'čuk 1974), although none of the points being made here hangs in any way on these notations. In particular, we will have recourse to the notion of a "morphological rule" and the corresponding formalisms, which are part of the morphological module of a Meaning-Text linguistic model. Obvious constraints of space prevent us from discussing Meaning-Text morphological models in general; we will limit ourselves to a few references: Mel'čuk (1967, 1976, 1993–2000: vol. 5, 2006a) and Beck et al. n.d.).

13. The lexicographic numbers used here refer to entries and subentries in the *Longman Dictionary of Contemporary English Online* (http://www.ldoceonline.com/).

14. The suffix -*mą* has two allomorphs, -*mą* and -*m*, the former following consonants and the latter following vowels.

15. The Ø- third-person object marker is treated as a prefix in (10) by analogy with the first-person object prefix because first-person singulars and third-persons generally have more in common with each other, morphologically speaking, than with second-persons. For example, the first-person singular and third-person plural subject markers are prefixes, as opposed to the suffixal second-person and first-plural markers; first-person singular and third-persons pattern together, in contrast to second person, with respect to aspectual forms; and first and third persons pattern together in the indefinite voice, being expressed by object prefixes, while the second-person is expressed by subject suffixes (Beck 2004).

16. It occurs to us that inflectional phrasemes could in fact exist at the phrasal level as well, although we are as yet unaware of any convincing examples. This situation might hold in a language that makes heavy use of periphrastic inflectional means such as case-marking particles, articles, and/or auxiliaries. Such a language might phraseologize a particular combination of analytical grammatical signs to express a particular inflectional meaning, making that combination into a restricted expression of a grammatical meaning.

17. Morphological phrasemes could, of course, also consist in combinations of segmental signs and nonconcatenative morphological operations; however, we will confine ourselves here to discussion only of affixal morphemes for ease of presentation.

18. In fact, we know of only one example of compounding used as an inflectional means (see Beck in press). If this pattern turns out to be more prevalent than previously supposed, it may be that the taxonomy of morphological phrasemes will require an additional category.

19. Plag (2002) argues that many restrictions of this type are not entirely idiosyncratic, but can be described as being driven by the phonological, semantic, and morphosyntactic properties of the affixes involved; even from this perspective, however, we would still consider the verb to be the base of the collocation: the verb is selected freely (based on meaning) and it is this verb against which the affix's combinatorial potential is measured. Thus, it is the verb that determines whether or not a particular affix is appropriate, given the affix's selectional restrictions.

20. This is, of course, an issue that is far too deep to go into here, but the crux of the matter resides in the fact that to treat a situation such as that in Burushaski as being a case of rule-governed allomorphy would require, in essence, the creation of a different "allomorphy" rule for each lexical root. This is, at best, the equivalent of having each form of the plural suffix specified in the lexical entry for that root, and seems at odds with the more usual type of allomorphic rule which generalizes over large sets of forms without recourse to the individual lexical entries of the bases involved. See Mel'čuk (1993–2000: vol. 4, 203–207) for more extensive discussion. The restriction of allomorphy to cases where "sufficiently general" rules can be posited should not be construed as ruling out suppletive allomorphy. It is entirely possible to provide for allowable exceptions in a system which is largely regular or rule-governed: thus, in English, we have the nominal plural suppletive allomorph -*en* (in *brethr–*

en, ox–en and *childr–en*), which is an exception with respect to the regular /s/ ~ /z/ ~ /ɪz/ allomorphy.

21. The alternation *-čaj* ~ *-če* is strictly morphophonological, having to do with the interaction of the ergative and the oblique case suffixes.

22. UNT verbs are inflected for four aspects — incompletive (action temporally unbounded), completive (action temporally bounded), progressive (action continuing at time of reference), and perfect (action completed in past such that its results are relevant at time of reference).

23. The fact that the second of the two affixal morphemes is indeed the completive marker (rather than an accidentally homophonous segment) is made evident by the allomorphy of the completive suffix, which is the same as that of the second part of the progressive marker in all contexts. For example, consider the examples in (i)–(iv), which illustrate the aspectual forms of the verb *musú:* 'kiss' in the indefinite voice:

    (i) *kimusu:kán*
        kin–musú:–kan–ya:
        1$_{OBJ}$–kiss–INDF–INCOMP
        'I am kissed' [Fr. *On m'embrasse;* here and in (ii)–(iv) below 'I' is Direct Object]

    (ii) *kimusu:kaní:*
        kin–musú:–kan–ni̠:
        1$_{OBJ}$–kiss–INDF–PERF
        'I have been kissed'

    (iii) *kimusú:ka̠*
        kin–musú:–ka̠
        1$_{OBJ}$–kiss–INDF:COMP
        'I got kissed'

    (iv) *kimusu:má:ka̠*
        kin–musú:–ma:–ka̠
        1$_{OBJ}$–kiss–PROG–INDF:COMP
        'I am being kissed'

In the incompletive and perfect aspects (i) and (ii), the marker of the indefinite voice is *-kan*; this suffix is easily identifiable and combines with the aspect- and person-markers in regular and predictable ways. In the completive aspect (iii), however, the indefinite voice is cumulatively expressed along with aspect by the suppletive portmanteau form *-ka̠*. This same affix is found in the indefinite progressive form (iv) as well, where it remains a demonstrably separate affix from *-ma:* (cf. *ka:ta:wama:ʔó:ka̠* 'they are all being given food (*ta:wá*)'). The relative position of the indefinite marker in the suffix string is also different for verbs in the incompletive and perfect aspects than it is in the completive and progressive aspects: in (i) and (ii), *-kan* appears before the aspect suffixes (position 4 in the morphological template — Beck et al. n.d.), whereas in (iii) and (iv) it appears much further from the stem (position 12), as can be seen by the fact that *-ka̠* follows *-ma:* (which appears in position 5, like the incompletive and perfect markers).

    The aspectual form of the verb also triggers alternation in the repetitive suffix *-pala*, the incompletive and perfect forms of the verb taking one allomorph (the full suffix, *-pala*) and the completive and progressive forms another (*-pa*):

    (v) *a̠mpalá*
        Ø–a̠n–pala–ya:
        3SG$_{SUB}$–go–RPT–INCOMP
        's/he goes again'

    (vi) *a̠mpalaní:*
        Ø–a̠n–pala–ní:

3SG_{SUB}–go–RPT–PERF
's/he has gone again'

(vii) *ampáł*

Ø–a̱n–pala–li̱

3SG_{SUB}–go–RPT–COMP

's/he went again'

(viii) *ama:páł*

Ø–a̱n–ma:–pala–li̱

3SG_{SUB}–go–PROG–RPT–COMP

's/he is going again'

These facts reflect the diachronic origin of the progressive marker (the completive form of a positional verb, *MA: 'be lying down'), but show the desirability of treating the second part of the progressive marker as a semantically empty instantiation of the completive morpheme.

24. In this case it is, of course, logically possible that matters be reversed and that [17b] be written as a rule for the expression of first-person objects when their number is plural:

(i) $1_{\text{OBJ}} \Leftrightarrow \{1_{\text{OBJ}}\} \oplus \{2_{\text{OBJ}}\} \mid \text{number}_{\text{OBJ}} = \text{PL}$

However, given that most of the special cases in UNT person-marking involve person-hierarchy effects (Beck 2001) and have to do with the expression of person rather than number (see, for example, the discussion of [32c] and [36d], and, especially, the rule in [35]), it seems more natural to describe this formation as conditioned by the person of the object rather than by its number. Including the person of the object as a condition on the rule, rather than writing it into the rule itself, is a technical choice made in the interests of maximalizing the generality of the rules in the system — in particular, the rule for the expression of the first-person object in [17a], which is excluded from applying to first-person plural objects by the existence of the more specific rule in [17b].

25. A notable exception here is Pike (1961), who describes a particular type of restricted complex of suffixes (what we discuss under the heading of "inflectional idioms" in Section 3.2.4) in Ocaina as an "affixal idiom." Our thanks to an anonymous reviewer for drawing our attention to this source.

26. The component given in SMALL CAPS in the gloss is, once again, the idiom's semantic pivot.

27. This is, of course, a different situation from the case of inflectional idioms, discussed below, where the conventionalized usage of affixes is systematic and maintaining the literal glosses of inflectional morphemes used idiomatically represents a considerable formal economy (see Note 23 and Section 3.2.4).

28. Etymologically speaking, this circumfix is part of the word *Lushootseed* itself, which is properly *dxʷləšutsid*, the root *ləš* being an old proto-Salish form meaning 'people' and forming part of the term "Salish" (T. Hess p.c.). The word is, however, unanalyzable in synchronic terms.

29. The middle voice suffix here is required because the lexical suffix *-us* 'face' expresses a body part of the agent/subject.

30. We would contend this is true even for markers of number of subject and object, which at first glance might seem to be predicates bearing on the person of the subject/object. Note that, for example, the meaning '3pl.sub' is not 'pl'('3sub'), which would be 'a plurality of subjects that are third-person': there is only one subject, that subject being the expression of a plurality of third-persons. This is, of course, a profound topic, but one that is well beyond the scope of this paper.

31. The third inflectional category, exclusivity, includes two grammemes, INCLUSIVE and EXCLUSIVE, and (in UNT) applies only to the first-person plural subject. Such systems are open to other interpretations, discussion of which is beyond the scope of this paper (see Cysouw 2003 and references therein).

32. This rule is a re-formulation of the rule given for heuristic purposes in (8). The necessity for the additional condition on its application follows from the discussion of the complementary rule in (35).

33. For instance, the realization of the past tense prefix *iš-* as *ša-* in the presence of the first-person singular subject prefix in (44), which applies in both its free and restricted uses.

34. Indeed, it is for this reason that it might be better to speak of inflectional "signifieds" rather than "meanings;" however, we have taken the second option in the interests of making our discussion more readable, in spite of the tension this has created between the meaning of "meaning" when used to describe lexical and derivational signifieds and the meaning of "meaning" when used to describe inflectional signifieds.

References

Aikhenvald, Alexandra Y. 2007. Typological distinctions in word formation. In Timothy Shopen (ed.), *Language typology and syntactic description, volume III: Grammatical categories and the Lexicon*, 2nd edn., 1–65. Cambridge: Cambridge University Press.

Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge & New York: Cambridge University Press.

Aronoff, Mark. 1994. *Morphology by itself: Stems and inflectional classes*. Cambridge, MA: MIT Press.

Aronoff, Mark & S. N. Sridhar. 1984. Agglutination and composition in Kannada verb morphology. In David Testen, Veena Mishra & Joseph Drogo (eds.), *Papers from the Parasession on Lexical Semantics*, 3–20. Chicago: Chicago Linguistic Society.

Baerman, Matthew, Greville G. Corbett, Dunstan Brown & Andrew Hippisley (eds.) 2007. *Deponency and morphological mismatches*. Oxford: Oxford University Press.

Bally, Charles. 1951 [1909]. Traité de stylistique française. [Treatise of French Stylistics.] Vol. 1. Genève: Georg/Paris: Klincksieck.

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

Beck, David. 2001. Person-hierarchies and the origin of asymmetries in Totonac verbal paradigms. *Linguistica Atlantica* 23. 1–33.

Beck, David. 2004. *Upper Necaxa Totonac*. Munich: Lincom Europa.

Beck, David. 2007. Morphological phrasemes in Totonacan inflection. In Kim Gerdes, Tilmann Reuther & Leo Wanner (eds.), *Meaning-text theory 2007: Proceedings of the 3rd International Conference on Meaning-Text Theory* (Wiener Slawistischer Almanach Sonderband 69), 107–116. Munich: Otto Sagner.

Beck, David. In press. Lexical, quasi-inflectional, and inflectional compounding in Upper Necaxa Totonac. In Alexandra Y. Aikhenvald & Pieter Muysken (eds.), *Multi-verb Constructions: A View from the Americas*. Leiden: Brill.

Beck, David, Joshua Holden & Vianey Varela. n.d. Upper Necaxa Totonac non-derivational morphology. Unpublished ms., University of Alberta.

Berger, Hermann. 1974. *Das Yasin-Burushaski* (*Werchikwar*). [Yasin-Burushaski (Werchikwar).] Wiesbaden: Otto Harrassowitz.

Blevins, James P. 2003. Stems and paradigms. *Language* 79. 737–767.

Croft, William A. 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Cysouw, Michael. 2003. *The paradigmatic structure of person marking*. Oxford: Oxford University Press.

Fabb, Nigel. 1988. English suffixation is constrained only by selectional restrictions. *Natural Language and Linguistic Theory* 6. 527–539.

Fillmore, Charles, Paul Kay & Catherine O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64. 501–38.

Goldberg, Adele. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Harley, Heidi & Elizabeth Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 73. 482–526.

Haspelmath, Martin. 2002. *Understanding morphology*. London: Arnold.

Hess, Thom. 1998. *Lushootseed reader with intermediate grammar, vol. 2*. Missoula, MT: University of Montana Occasional Papers in Linguistics.

Hockett, Charles. 1966. What Algonquian is really like. *International Journal of American Linguistics* 32. 59–73.

Holden, Joshua. 2009. Towards a new meta-language for Athapaskan linguistics: The case of morphological phrasemes. In David Beck, Kim Gerdes, Jasmina Milićević & Alain Polguere (eds.), *Proceedings of the Fourth International Conference on Meaning-Text Theory*, 157–165. Montreal: University of Montreal.

Kibrik, Aleksandr. E. 1997. Ierarxii, roli, nuli, markirovannost' i "anomal'naja" upakovka grammatičeskoj semantiki. [Hierarchies, roles, zeroes, markedness and "anomalous" wrapping of grammatical semantics.] *Voprosy jazykoznanija* 4. 27–57.

Kiparsky, Paul. 1982. From cyclic phonology to lexical phonology. In Harry van der Hulst & Norval Smith (eds.), *The structure of phonological representations, vol. I*, 131–175. Dordrecht: Foris.

Jackendoff, Ray. 2002. *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.

Langacker, Ronald. 1987. *Foundations of cognitive grammar*, vol. 1. Stanford: Stanford University Press.

Langacker, Ronald. 1991. *Foundations of cognitive grammar*, vol. 2. Stanford: Stanford University Press.

Matthews, Peter H. 1972. *Inflectional morphology: A theoretical study based on aspects of Latin verb conjugation.* Cambridge: Cambridge University Press.

Matthews, Peter H. 1991. *Morphology*. Cambridge: Cambridge University Press.

Mel'čuk, Igor A. 1964. OboObščenie ponjatija frazeologizma (morfologičeskie "frazeologizmy"). [Generalizing the notion of phraseologism (morphological "phraseologisms").] In Leonid I. Rojzenzon (ed.), *Materialy konferencii "Aktual'nye voprosy sovremennogo jazykoznanija i lingvističeskoe nasledie E.D. Polivanova"* [Proceedings of the conference "Current Problems in Modern Linguistics and the Linguistic Heritage of E. D. Polivanov."], 89–90. Samarkand: Samarkandskij Gosurdarstvennyj Universitet.

Mel'čuk, Igor A. 1967. Model' spraženija v ispanskom jazyke. [A model of Spanish verb conjugation.] *Mašinnyj perevod i prikladnaja lingvistika* 10. 21–53.

Mel'čuk, Igor A. 1974. *Opyt teorii lingvističeskix modelej "Smysl ⇔ Tekst"*. [An outline of the theory of "meaning ⇔ text" type linguistic models.] Moscow: Nauka.

Mel'čuk, Igor A. 1976. A model of Spanish verb conjugation. In Igor A. Mel'čuk, *Das Wort: Zwischen Ausdruck und Bedeutung*, 210–257. Munich: Wilhelm Fink.

Mel'čuk, Igor A. 1982. *Towards a language of linguistics: A system of formal notions for theoretical morphology*. Munich: Wilhelm Fink.

Mel'čuk, Igor A. 1993–2000. *Cours de morphologie générale et descriptive*. [A course in general and descriptive morphology.] Montreal: Les Presses de l'Université de Montréal.

Mel'čuk, Igor A. 1995. Phrasemes in language and phraseology in linguistics. In Martin Everaert, Eric-Jan van der Linden, André Schenk & Rob Schreuder (eds.), *Idioms: Structural and psychological perspectives*, 167–232. Hillsdale, NJ: Lawrence Erlbaum.

Mel'čuk, Igor A. 2004. La non-compositionnalité en morphologie linguistique. [Non-compositionality in linguistic morphology.] *Verbum* 26. 439–458.

Mel'čuk, Igor A. 2006a. *Aspects of the theory of morphology*. Berlin & New York: Mouton de Gruyter.

Mel'čuk, Igor A. 2006b. Explanatory combinatorial dictionary. In Giandomenico Sica (ed.), *Open problems in linguistics and lexicography*, 222–355. Monza: Polimetrica.

Mel'čuk, Igor A. 2008. Zero affixes and nominal cases in Daghestanian languages. In Aleksandr Arxipov, Leonid Zaxarov, Andrej A. Kibrik, Aleksandr E. Kibrik, Irina Kobozeva, Ol'ga Krivnova, Ekaterina Ljutikova & Ol'ga Fëdorova (eds.), *Fonetika i nefonetika. K 70-letiju Sandro V. Kodzasova* [Phonetics and non-phonetics. For the 70th birthday of Sandro V. Kodzasov.], 176–183. Moscow: Jazyki Slavjanskix Kul'tur.

Mithun, Marianne. 1984. The evolution of noun incorporation. *Language* 60. 847–895.

Pike, Kenneth L. 1961. Compound affixes in Ocaina. *Language* 37. 570–581.

Plag, Ingo. 2002. The role of selectional restrictions, phonotactics and parsing in constraining suffix ordering in English. In Geert Booij & Jap van Marle (eds.), *Yearbook of morphology 2001*, 285–314. Dordrecht: Kluwer.

Schultze-Berndt, Eva. 2000. *Simple and complex verbs in Jaminjung. A study of event categorization in an Australian language* (MPI Series in Psycholinguistics 14). Nijmegen: Max Planck Institute for Psycholinguistics.

Seifart, Frank. 2005. *The structure and use of shape-based noun classes in Miraña (North West Amazon)* (MPI Series in Psycholinguistics 33). Nijmegen: Max Planck Institute for Psycholinguistics.

Spencer, Andrew. 2001. The paradigm-based model of morphosyntax. *Transactions of the Philological Society* 99. 279–313.

Stump, Gregory T. 2001. *Inflectional morphology: A theory of paradigm structure*. Cambridge: Cambridge University Press.

Švedova, Natal'ja Ju. (ed.). 1980. *Russkaja grammatika, v. I.* [Russian Grammar, vol. I.] Moscow: Nauka.

Žolkovskij, Aleksandr K. & Igor' A. Mel'čuk. 1965. O vozmožnom metode i instrumentax semantičeskogo sinteza. [On a possible method and instruments for semantic synthesis.] *Naučno-texničeskaja informacija* 5. 23–28.

Žolkovskij, Aleksandr K. & Igor' A. Mel'čuk. 1967. O semantičeskom sinteze. [On semantic synthesis.] *Problemy kibernetiki* 19. 177–238.

Zwicky, Arnold M. 1985. How to describe inflection. *Berkeley Linguistics Society* 11. 371–386.

Zwicky, Arnold M. 1992. Some choices in the theory of morphology. In Robert D. Levine (ed.), *Formal grammar: Theory and implementation*, 327–371. Oxford: Oxford University Press.