PHRASEOLOGY IN MORPHOLOGY: IT'S A SIGN![1]

David Beck
University of Alberta

ABSTRACT

A well-studied property of the linguistic sign at the level of multi-word expressions is the capacity for these expressions to take on conventionalized and context-specific meanings, resulting in what are called phrasemes (Mel'čuk 1995). What is less well-studied is the existence of restricted expressions at the level of the word, what would be called morphological phrasemes. Consideration of restricted morphological expressions reveals that these, on the whole, fall into the same types as multi-word phrasemes. This paper shows that this parallelism falls out directly from the nature of the morph as a linguistic sign and the susceptibility of linguistic signs at all levels of expression to the processes of phraseology. Recognizing this allows us to resolve a number of issues that have been used to argue against the morpheme as an analytical unit, and allows us to account for certain kinds of problematic constructions found in natural language.

Keywords: phraseology, morphology, morphological idiom, morphological collocation, idiom, phraseme.

1. COMPOSITION IN MORPHOLOGY

A common assumption made about morphologically complex expressions, at least in the realm of inflection, is that they must be compositional – that is, that the meaning of the base plus the in-

[1] I would like to thank, first and foremost, Igor Mel'čuk, whose work has been the inspiration for the ideas that underlie this paper and whose collaboration on this topic has made the paper possible in the first place. I would also like to thank Jasmina Milićević and the audience at the 42nd Annual Meeting of the Atlantic Provinces Linguistic Association, as well as the editor and a reviewer for Linguistica Atlantica, for helpful and positive feedback on this paper. Anything that is still wrong with it is my responsibility alone.

BDD-A30748 © 2019 Atlantic Provinces Linguistic Association
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-15 01:55:00 UTC)

flectional affixes should be the sum of the meanings of parts that make it up. However, data like these from Upper Necaxa Totonac (UNT) are problematic for this view:[2]

(1) i̱kłtatá:
i̱k–łtata:–Ø
1sg.sub–sleep–impf
'I sleep'

(2) łtata:yá:ʍ
łtata:–ya:–ʍ
sleep–impf–1pl.sub:incl
'we-including-you sleep'

(3) i̱kłtata:yá:ʍ
i̱k–łtata:–ya:–ʍ
1sg.sub–sleep–impf–1pl.sub:incl
'we-excluding-you sleep'

In the examples in (1) and (2), we see straightforward compositional forms in which the prefix *i̱k-* signifies 'first person singular subject' and the suffix *−w* signifies 'first person plural inclusive subject'; however, in (3) we see that the combination of these same affixes gives the reading 'first person plural exclusive subject', contradicting both the 'singular' part of 'first person singular subject' and the 'inclusive' part of 'first person plural inclusive subject'. Thus, the form in (3) is non-compositional ('I' ⊕ 'we including you' ≠ 'we excluding you'). At the lexical level, such a non-compositional combination of elements – that is, words that combine together to mean something unexpected – would be described as an idiom, suggesting that the combination of affixes in (3) might be described as a *morphological idiom*.

The term "idiom" is not often used at the morphological level, largely because it is assumed that morphological expressions, unlike lexical expressions, are not subject to the processes of conventionalization and phraseology that create idioms and other set or non-free expressions[3], often referred to collectively as *phrasemes*. There is, however, some work (Mel'čuk 2004; Beck 2007; Beck & Mel'čuk 2011) that calls this assumption into question, taking as its starting point the simple notion that morphemes are in themselves linguistic signs and that phraseology is applicable to signs at both levels, lexical and morphological. Beck & Mel'čuk (2011) elaborate in detail the parallels between phrasemes of all types at the two levels and assert that these parallels

2 Upper Necaxa Totonac [iso tku] is a Totonacan language spoken by 3,000 people in the Sierra Norte of the state of Puebla, Mexico. All uncited Totonac data in this paper are from the author's fieldnotes. The abbreviations used in examples are as follows: 1, 2, 3 first-, second, third-person; BEN benefactive; EXCL exclusive; IMPF imperfective; INCL inclusive; IND indicative; OBJ object; PFV perfective; PL plural; PROG progressive; RECIP reciprocal; SG singular; SUB subject; TOT totalitive. Totonac examples are given in an Americanist Phonetic Alphabet; the acute accent in transcriptions represents word-level stress. Examples from other languages follow common orthographic conventions for those languages or the conventions of their sources.

3 Although P. De Decker (p.c.) notes that phraseologization is an acknowledged principle of diachronic word formation, even if the term itself is not widely applied.

BDD-A30748 © 2019 Atlantic Provinces Linguistic Association
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-15 01:55:00 UTC)

constitute strong evidence that morphemes are indeed linguistic signs *par excellence*. At the end of their paper, Beck & Mel'čuk note that treating morphemes as elementary signs seems to be at odds with some current approaches to morphology that advocate against linguistic representation of meaning below the level of the word (e.g., Anderson 1992; Haspelmath 2003; Blevins 2006), and it is this idea I would like to take up here, with an eye towards showing, ultimately, that there is evidence both from phraseology and from diachrony that speakers do treat sublexical units in some of the ways that they treat lexical signs. To that end, this paper begins with a review of the concept of the linguistic sign and Mel'čuk's (1982) typology of phrasal phrasemes (§2). In section 3, I show the same typology applied to morphological phrasemes, recapitulating for the most part (and in less detail) the discussion in Beck & Mel'čuk (2011). Section 4 examines the tension mentioned above between the view of the morpheme as a linguistic sign and a whole-word approach to morphology which advocates against analysis of signs smaller than the word. Although we see in §4 that there is indeed good reason to believe that speakers are aware of and can assign meanings to sublexical units, in the conclusion to the paper (§5), I suggest that, in fact, the apparent conflict between analytical and analogical approaches to morphological representations is more an artifact of the metaphors used by linguists than it is of the nature of language itself.

2. LINGUISTIC SIGNS

The lexicon of a language is generally taken to be an inventory of the linguistic signs available to speakers. Although there are many modern variations on (and terms for) linguistic signs, all coincide in that signs are conceived of as conventionalized mappings between meaning and form – or, in the classic formulation of the sign, signified (signifié) and signifier (signifiant) (de Saussure et al. 2011). In a more comprehensive approach to the lexicon, it is also necessary to add to this pair a consideration of the combinatorial and distributional properties of signs, what Mel'čuk (1982, inter alia) refers to as their "syntactics". Thus, a linguistic sign consists of a triplet – signified, signifier, and syntactics, as shown in Figure 1:

FIGURE 1

The tripartite linguistic sign

‘domesticated carnivorous mammal of the species *Canis familiaris*’	/dɑg/	{noun}

Such elementary signs consisting of a singleton, irreducible unit corresponding to a single fairly concrete meaning are the canonical type of sign; however, it is by no means the case that all linguistic signs are this straightforward. The fact is that a linguistic sign can originate in any conventionalized (usage-based) association of sound (and syntactics) to meaning, and these associa-

BDD-A30748 © 2019 Atlantic Provinces Linguistic Association
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-15 01:55:00 UTC)

tions are well known to hold both above and below the level of the word. Signs smaller than a word are morphs, and these can be combined to form complex multi-morphemic signs. Likewise, multi-word expressions often take on conventionalized meanings particular to certain linguistic or pragmatic contexts. Such expressions constitute a class of signs referred to as *phrasemes* (Mel'čuk 1982, 1995).

Phrasemes, under other names, are well-studied (e.g., Bally 1951; Everaert et al. 1995), and the most systematic taxonomy of these expressions is that of Mel'čuk (1982), where they are divided into types based in the first place on their compositionality – that is, whether the meaning of the expression as a whole is the sum of the meaning of its parts. *Compositional phrasemes* are those expressions whose meaning is in this sense compositional but whose parts, unlike those of non-conventionalized or "free" expressions, are not all freely chosen by the speaker, being restricted in some way by context. Mel'čuk identifies three sub-types of compositional phraseme – pragmatemes, clichés, and collocations.

Pragmatemes are set phrases that are required by convention in specified circumstances such as those shown in (4):

(4) Eng. *Best before* [on food for sale]; cf. Rus. *Srok godnosti* 'deadline of fitness'
Eng. *No smoking* [on sign]; cf. Span. *No fumar* 'not to smoke', Fren. défense de fumer 'interdiction of smoking'

Each of these expressions is compositional in the sense that the intended meaning is a predictable meaning of the sum of its constituent elements, but each is prescribed for a particular set of circumstances (shown in the square brackets) and none of the several synonymous expressions – such as the translation equivalents shown for other languages – would be used by speakers under those conditions.

Clichés, on the other hand, are less restricted in that they are substitutable by free expressions with the same meaning, but they are largely expected and relied upon to express certain meanings in typical circumstances:

(5) Eng. *You've seen one you've seen 'em all.*
Eng. *in the wrong place at the wrong time*

Clichés are typically valued not just for their literal meanings, but for the connotations that go with them: *you've seen one you've seen them all* conveys not just the homogeneity of the objects of description but also a certain disdain, while *in the wrong place at the wrong time* carries with it the implication that some harm or misfortunate occurred to the person being talked about.

The final type of compositional phraseme, the *collocation*, consists of (at least) two parts, one of which, the base (in small caps in (6) below) is selected freely, for its meaning, while the other part, the *collocate*, is selected as a function of that base:

(6) Eng. *heavy rain, sound argument, pay attention, crows caw*

BDD-A30748 © 2019 Atlantic Provinces Linguistic Association
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-15 01:55:00 UTC)

For an English speaker, an intense rainfall is a *heavy rain* (cf. *strong wind*, #*strong rain*) and the use of *heavy* to express intensity is restricted to this and a handful of other bases (*heavy snow*, *heavy buffeting*; #*heavy wind*). *Sound* meaning 'good, rigorous' is selected for by a few nouns denoting reasoning (*sound [mathematical] proof*, *sound logic*, etc.). These two collocations do co-exist with compositional equivalents, but the only verb that goes with *attention* is *pay* and the only correct way to express the sound made by a *crow* is *caw*[4].

These three types of compositional phraseme are contrasted with the *non-compositional phraseme* or *idiom*, which Mel'čuk defines as an expression whose meaning is not the meanings of the sum of its parts and none of whose components are chosen freely. These also fall into three categories:

Weak idiom: an idiom that includes the meaning of all its components but none of whose components expresses its semantic pivot (see below)

> ⌜lightning rod⌝ 'device that is designed to protect constructions against ***lightning*** and has the form of a conducting ***rod***, fixed …'

Semi-idiom: an idiom that includes the meaning of only one of its components and none of whose components expresses its semantic pivot

> ⌜sea dog⌝ 'person having significant experience of navigation on the ***seas***'

Strong idiom: an idiom that does not include the meanings of any of its components but none of whose components expresses its semantic pivot

> ⌜shoot the breeze⌝ 'chat leisurely'

The three types of idiom are differentiated based on the degree to which their components literally express elements of their conventionalized meanings. In a weak idiom, the meanings of all the components, shown in bold italic in the glosses above, are part of the meaning of the expression – a ⌜lightning rod⌝ is shaped like a ***rod*** and is used to deal with ***lightning***. A semi-idiom has the meaning of only one component in its definition (⌜sea dog⌝ is related to the ***sea*** but has nothing to do with dogs), whereas in a strong idiom there is no relationship between the words that make up the expression and what it actually means.

An important component of the definitions of all of these types of idiom is the requirement that none of their components (the words that make them up) express their ***semantic pivot***, which can be defined as

> For an expression AB having the meaning 'S' such that 'S' = 'S1' ⊕ 'S2', the semantic pivot of 'AB' is that part 'S1' of AB's meaning 'S' such that 'S' [= 'S1' ⊕ 'S2'] can be represented as a predicate 'S2' bearing on 'S1', i.e., 'S' = 'S2'('S1') (Mel'čuk 2006: 277)

[4] It should be noted that the issue of compositionality in collocations is somewhat more complex than it is in the case of pragmatemes and clichés. The reader is referred to Mel'čuk (1995) for discussion.

BDD-A30748 © 2019 Atlantic Provinces Linguistic Association
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-15 01:55:00 UTC)

In the examples above, the semantic pivot is given in small caps in the definition and, informally, represents the basic type to which the referent of the idiom belongs. Thus, a ˹lightning rod˺ is a type of device, a ˹sea dog˺ is a kind of person, and ˹shoot the breeze˺ is a type of chatting. The remainder of the idiom's meaning serves to qualify or specify the semantic pivot, which is itself not expressed by any component of the idiom itself. The notion of semantic pivot is important for us here both because it allows us to distinguish idioms from collocations (which do express their semantic pivots) and because the notion of the semantic pivot can be appealed to to account for an important property of non-compositional phrasemes in the realm of morphology (§3.2.2).

3. MORPHOLOGICAL PHRASEMES

While phrasemes are universally recognized and accepted as a class of multi-word linguistic signs, what is less widely recognized is that the same processes that lead to the formation of phrasal phrasemes potentially applies at the level below the word. Relatively little has been said about this in the literature, although conventionalized morphological expressions have occasionally been recognized (e.g., Aronoff & Sridhar 1984), and Pike (1961) uses the term "idiom" to describe conventionalized morphological forms in Ocaina. More recent work, however, has specifically argued for analyzing such expressions as morphological phrasemes (Mel'čuk 2004; Beck 2007; Beck & Mel'čuk 2011). This work has shown that, on the whole, morphological phrasemes fall into the same types as phrasal phrasemes (Figure 2), with certain restrictions and qualifications that fall out from the different nature of morphological, as opposed to lexical, signifiers. In the remainder of this section, I will review these parallels, and some of the divergences, for each of the attested types of morphological phraseme.

FIGURE 2

Types of phrasal and morphological phrasemes (Beck & Mel'čuk 2011: 215)

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

✓ = attested; ✗ = not possible; ? = as yet unattested; full = base + affixes; affixal = affixes only

BDD-A30748 © 2019 Atlantic Provinces Linguistic Association
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-15 01:55:00 UTC)

2.1. Compositional morphological phrasemes

Compositional phrasemes are those expressions that are restricted by pragmatic or linguistic context but which are compositional in that their meanings are the sums of the meanings of their parts. Compositional phrasemes fall into three types – pragmatemes, clichés, and collocations. Of these three, only collocations are common on the morphological level, so we will begin with these before considering some of the reasons that morphological pragmatemes and clichés might be difficult, or perhaps impossible, to find.

2.1.1. Morphological collocations

Collocations are those constructions in which one member, the base, is selected freely for its meaning while the other is restricted, being chosen for its base. We see an English example at the phrasal level in (7):

(7) Eng. *high fever (*hot fever, *strong fever*, etc.)

In this expression, the base, FEVER, is selected freely (the speaker is talking about a sickness involving elevated body temperature) whereas the intensifier, HIGH, is selected by the base, *high fever* being the most natural (perhaps only natural) way of referring to upper extremes of body temperature using the word *fever*. Compare (7) to the Russian collocational compound *konevodstvo* 'horse breeding' in (8):

(8) Rus. *kon–e–vodstv*-(*o*) 'horse breeding' (KON′ 'horse' + VODSTV 'breeding')
[cf. **lošad–e–vodstv*-(*o*) (LOŠAD′ 'horse' ≈ KON′ 'horse')]
(Beck and Me'čuk 2011: 191)

Here we have a base, VODSTV 'breeding', which is selected freely, and a collocate KON 'horse', which is selected by the base. The more common word for 'horse' in Russian is LOŠAD′, but it is impossible to use this word in a compound with VODSTV.

Finding a parallel situation in the realm of derivational morphology is easy, although we are forced to introduce a further distinction to our taxonomy, that between a *full morphological phraseme* and an *affixal phraseme*. A full phraseme is one in which the base is the root or stem to which the affixation applies, making the word as a whole a morphological collocation. Full derivational phrasemes are actually quite familiar and arise in cases such as the English inhabitant suffixes shown in (9):

(9) Eng. Montreal-er (*Montreal-ian), Edmonton-ian (*Edmonton-er), etc.

For words like these, the base (*Montreal*, *Edmonton*) is selected freely while the particular suffix is selected by the base, turning the entire word into a full morphological collocation.

BDD-A30748 © 2019 Atlantic Provinces Linguistic Association
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-15 01:55:00 UTC)

An affixal collocation does not involve the root or stem or the word, but instead occurs when the presence of one affix, selected for its meaning, requires the presence of another. This is the case of Upper Necaxa Totonac causative forms, shown in (10):

(10) UNT *ma:–łi:–ní:* 'cause (*ma:*-) someone to dance (*łi:*)'
[cf. causative of stative verbs, *ma:–tsumá:* 'cause something to be full (*tsumá:*)']

The base of the collocation is *ma:*- which conveys the desired meaning 'causative' and the suffix *-ni:* is required by the base. While it might be possible to attribute the meaning 'causative' to both parts of the collocation (treating them as a regular circumfix), the use of the *ma:*- prefix in other contexts on its own with the same meaning suggests that *-ni:* is, at least synchronically, semantically empty in these constructions. Given that *ma:*- 'causative' ⊕ *-ni:* 'Ø' ≡ *ma:*- *-ni:* 'causative', the expression qualifies as compositional and falls under the heading of an affixal collocation.

As with derivation, the realm of inflection also provides us with examples of both full and affixal collocations. An example of the first is provided by the Russian aspectual prefixes shown in Table 1 and in (11):

(11) Rus. *pro*- 'perfective' + *čitat′* 'read' [**vyčitat′* 'read']
Rus. *vy*- 'perfective' + *pit′* 'drink' [≠ *propit′* 'spend money on drink']

For such wordforms, the base is the verbal radical, which is selected freely for its meaning, but the prefix is specific to the base. The use of an incorrect prefix is either ungrammatical (**vyčitat′*) or creates a different lexeme (e.g., *propit′* 'spend money on drink' – which is, in fact, a derivational semi-idiom, see 3.2.1).

TABLE 1

Russian verbal aspectual prefixes (Švedova 1980: 587–8)

	imperfective	perfective
read'	*čitat′*	*pro–čitat′*
wake'	*budit′*	*raz–budit′*
do'	*delat′*	*s–delat′*
build'	*stroit′*	*po–stroit′*
drink'	*pit′*	*vy–pit′*
make drink'	*poit′*	*na–poit′*

BDD-A30748 © 2019 Atlantic Provinces Linguistic Association
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-15 01:55:00 UTC)

An example of an affixal inflectional collocation is provided by the Upper Necaxa Totonac progressive aspect, which is signified by the combination of two elements, the suffix *-ma:* and a semantically empty use of the perfective *-l̨i̱*. Consider the forms in (12) – (14):

(12) pasá:ɬ
pasa:–l̨i̱
burn–pfv
'it burns$_{pfv}$ up'

(13) pasa:má:ɬ
pasa:–ma:–l̨i̱
burn–prog–pfv
'it is burning'

(14) pasa:ma:ʔó:ɬ
pasá:–ma:–ʔo:–l̨i̱
burn–prog–tot–pfv
'it is burning completely' [≈ 'it is completely on fire']

(12) shows the compositional use of *-l̨i̱* as a perfective aspect marker; the same affix (as shown by a range of morphological properties it exhibits, such as morphophonological reduction to *−ɬ*) also appears in the progressive form of the same verb in (13). The fact that this affix is still analyzable as separate from the progressive marker *-ma:* itself is shown in (14), where the two are separated by the totalitive suffix *-ʔo̱:*. Constructions such as this where the signifier of an inflectional category is obligatorily accompanied by a semantically empty partner are discussed elsewhere under the heading of *parasitic formations* (Matthews 1972; Aronoff 1994).

3.1.2. Morphological pragmatemes and clichés

Once we begin thinking in terms of phrasemes, morphological collocations turn out to be frequent and commonplace. Not so morphological pragmatemes and clichés. Whereas phrasal pragmatemes – compositional phrases selected by contextual restrictions, in (4) above – and clichés – compositional phrases selected by conventional usage (5) – are quite numerous, on the morphological level both pragmatemes and clichés continue to be unattested. The likeliest reason for this is the very different degree of choice a speaker enjoys in terms of selection from a wide range of lexical resources versus the generally very limited number of choices offered by morphology. Pragmatemes and clichés are created when one particular phrase out of range of possible expressions with the same meaning becomes conventionalized as the only or the most expected expression of that meaning for a given context. It is seldom the case in a language that there is a wide range of morphological expressions that have the same meaning. The morphological resources of the language are scant: synonymy and near synonymy, lexically commonplace, are rare in morphology. This makes morphological pragmatemes and clichés less likely to exist, though such things may eventually be found in languages with exceptionally rich morphological inventories.

BDD-A30748 © 2019 Atlantic Provinces Linguistic Association
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-15 01:55:00 UTC)

3.2. Morphological idioms

Non-compositional phrasemes, or idioms, are those expressions whose meanings are not the proper sums of the meanings of their constituent elements. At the phrasal level, they are referred to as *idioms*, suggesting that non-compositional morphological phrasemes merit the title *morphological idioms*. Like morphological collocations, morphological idioms are reasonably well-attested once we start looking for them, and occur on all three levels of morphological expression – compounding, derivation, and inflection. On the levels of compounding and derivation (that is, in the realm of word-formation), we see the same three types of idiom – weak, semi-, and strong – that we found on the phrasal level; inflectional idioms, however, are different and can only be strong idioms. In the sections below, we'll begin with the clearer cases, morphological idioms in word formation, before moving on to consider inflectional idioms and why these might be different.

3.2.1. Morphological idioms in word formation

As noted above, non-compositional phrasemes, or idioms, can be grouped into three types depending on to what extent the meanings of the components of the phraseme are part of the meaning of the idiom itself. In weak idioms, the meaning of the phraseme includes the meaning of all its components (shown in bold italics in the glosses below), but none of these components expresses the semantic pivot (small caps), as in the weak phrasal idiom in (15):

(15) Eng. ⌜salt shaker⌝ 'dispenser from which ***salt*** can be ***shaken***'

Weakly idiomatic compounds, like this one from Upper Necaxa Totonac, follow precisely the same pattern:

(16) UNT *šwaya:skɬ́:tɬ̱* '***fish-eating*** bird [=kingfisher]' [< *šwáyᵾ̱* 'eater of' + *skɬ́:tɬ̱* 'fish']

The name for the kingfisher in UNT is a compound composed of the deverbal noun *šwáyᵾ̱* 'eater of' and the word *skɬ́:tɬ̱* 'fish', both of whose meanings are part of the meaning of the compound itself, 'fish-eating bird'. Because the compound refers to a species of bird, bird is its semantic pivot, which is not expressed by either of its constituent elements. (Note that this is also what makes the compound non-compositional in that *šwaya:skɬ́:tɬ̱* does not refer to everything that eats fish, only to a particular type of bird that has been conventionally associated with the term 'eater of fish'.)

Weak idioms are also readily identifiable in the realm of derivation. A full derivational weak idiom from Upper Necaxa Totonac is shown in (17):

(17) UNT *loŋót* 'disease causing shivering as if the ***result*** of ***feeling cold*** [= malaria]'
[< *loŋ-* 'feel cold' + *-ot* 'result']

The Upper Necaxa word for malaria, *loŋót*, is a figurative expression formed from the resultative nominalizer (*-ot*) and the radical *loŋ-* 'feel cold'; the meaning of both of these components is in-

BDD-A30748 © 2019 Atlantic Provinces Linguistic Association
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-15 01:55:00 UTC)

cluded in the figurative meaning of the word, but neither directly expresses the notion of 'disease', the semantic pivot of the expression.

An affixal derivational weak idiom is the UNT circumfix *puː- -n* 'container-like instrument':

(18) UNT *puː-* V *-n* '***container-like*** tool used ***in an action*** V' [< *puː-* 'container' + *-n* 'thing involved in an action']
puːčapán 'mill, grinder' [< *čapá* 'grind sth']

The circumfix here contains the meanings of both of its component parts, although the abstract nature of the meaning of *-n* 'thing involved in an action' somewhat obscures its role in the definition as realized by the English gloss.

The next type of idiom, the semi-idiom, includes the meaning of only one of its components, as in the English phrasal idiom in (19):

(19) ˹eager beaver˺ 'person who takes on tasks or responsibilities ***eagerly***'

Here we see a phrasal expression denoting a type of person whose meaning includes that of one of its components, *eager*, but not the other, *beaver*. This is paralleled exactly by the compound semi-idiom in (20):

(20) Eng. *lumberjack* 'labourer who fells trees for ***lumber***'

Just as the phrasal semi-idiom in (19) does not refer to a 'beaver', neither does the compound word in (20) refer to a 'jack' (in any synchronic sense of the term). It does, however, contain the meaning of the word *lumber* in its semantics, making this a compound semi-idiom.

An example of a derivational full semi-idiom can be seen in (21):

(21) Eng. *locker* 'compartment, usually made out of metal, that can be ***locked***'

The word *locker* refers to a kind of compartment but does not express this meaning by either the root or the *-er* affix; the root does, however, express an important component of the word's meaning in that a locker isn't a locker if it can't potentially be locked.

Examples of derivational affixal semi-idioms are a little harder to come by, but the Russian circumfix *za- -j* seems to fit the bill:

(22) Rus. *za-* N *-j-* 'region located ***beyond*** (*za-*) natural obstacle N' [< *za-* 'beyond' + *-j*]
Zabajkal'j-(*e*) 'region east of [≈ beyond] Lake Baikal'
Zakarpat'j-(*e*) 'region east of [≈ beyond] the Carpathian mountains'
Zavolž'j-(*e*) 'region east of [≈ beyond] the Volga river'
(Beck & Mel'čuk 2011: 191)

Here we see that the circumfix's meaning includes that of *za-* 'beyond, on the other side of' but does not include the meaning of *-j*, which in fact has no clear meaning in this expression. As noted in Beck & Mel'čuk (2011), there are different ways of treating this situation, but all of them end up with the analysis of (22) as as semi-idiomatic affixal idiom.

BDD-A30748 © 2019 Atlantic Provinces Linguistic Association
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-15 01:55:00 UTC)

Strong idioms are those in which the meaning of none of their components are part of the meaning of the expression as a whole, as in the English phrasal idiom in (23):

(23) Eng. ˹bite the dust˺ 'die as the result of an attack'

Compounds that fit this pattern are widely recognized in the literature and extremely common. An example would be the English compound word in (24):

(24) Eng. *skinflint* 'person who is reluctant to spend money'

Skinflint refers to a person and designates them as a miser, but neither SKIN nor FLINT relate to any aspect of its meaning.

A full affixal strong idiom is illustrated by the Upper Necaxa word for 'caterpillar' in (25):

(25) UNT *kɻ́nɻ́* 'larva of a lepidopteran [= caterpillar]' [< *kɻ́n* 'swell' + *-nɻ́* 'nominalizer']

Once again, neither of the components of the word makes a direct contribution to its meaning, although the metaphor behind the formation is clear enough – like many idiomatic phrases and compounds, this expression is transparent in the sense of being easily psychologically accessible, but it is not compositional in the sense of being the proper sum of the meanings of its parts.

A derivational affixal strong idiom is illustrated in (26):

(26) Lushootseed *dxʷ-* V *-igʷəd* 'frame of mind' [< *dxʷ-* 'contained' + *-igʷəd* 'interior of body']
dxʷx̌ʷal'igʷəd 'give up, resign oneself to defeat' [< *x̌ʷal'* 'be helpless']
(Hess 1998: 18–19)

The Lushootseed circumfix consists of two parts, *dxʷ-* 'contained' and *-igʷəd* 'interior of body', which taken together express a frame of mind or emotional state; this circumfix is added to a moderate number of verbal bases to derive expressions of typical mental states associated with that verb.

3.2.2. Non-compositional inflectional phrasemes (inflectional idioms)

Although the parallelism between phrasal and morphological idioms holds for compounding and derivation, it breaks down somewhat at the level of inflection, where it turns out that the only possible morphological idioms are strong ones. This follows from the specialized nature of inflection, and the reasons for it become apparent when we compare strong phrasal idioms with strong inflectional idioms in terms of the correspondences between their meanings (semantics) and their morphological expression. Consider the regular, compositional use of the English lexeme BREEZE in the expression *cold breeze*:

(27) *cold **breeze*** | 'cold' ⊕ 'breeze' ⇔ {cold} ⊕ {breeze}

The expression in (27) is a free expression represented at the semantic level by the combination of two meanings, 'cold' and 'breeze', each of which maps on to lexemes on the morphological

BDD-A30748 © 2019 Atlantic Provinces Linguistic Association
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-15 01:55:00 UTC)

level. In this case, the meanings expressed by the words are their typical meanings and there is complete compositionality in their combination. Contrast this with the phrasal idiom in (28):

(28) ⌜shoot the breeze⌝ | ‘chat’ ⊕ ‘leisurely’ ⇔ {shoot} ⊕ {breeze}

Here we see that *shoot the breeze* expresses two meanings, ‘chat’ and ‘leisurely’, but neither of these meanings is the meaning of the components of the morphological expression, {shoot} and {breeze}. It is this non-compositionality, or mismatch between meaning and form, that is the hallmark of idioms, and the fact that there is a complete disjuncture between the meanings expressed and the forms used to express them is what determines this to be a strong idiom.

We can now compare the situation in (28) to the situation in (29) – (31), which present some inflected forms in Upper Necaxa Totonac. The first two forms, (29) and (30), show the formation of the indicative mood of the perfective aspect and the past tense, respectively:

(39) *pasa:ɬ* | ‘burn’ ⊕ ‘perfective’ ⇔ {burn} ⊕ {pfv}
Ø–pasa:–li̱
ind–burn–pfv
‘it burns up’

(30) *i̱špasá:* | ‘burn’ ⊕ ‘past’ ⇔ {burn} ⊕ {past}
Ø–i̱š–pasa:–Ø
ind–past–burn–impf
‘it burned’

In both these cases, we see compositional expressions of the two inflectional categories, the meaning ‘perfective’ being expressed by *-li̱* and the meaning ‘past’ expressed by *i̱š-*. When combined in a single verb form, however, the two affixes express a completely different meaning, the present tense of the irrealis mood:

(31) *i̱špasá:ɬ* | ‘burn’ ⊕ ‘present’ ⊕ ‘irrealis’ ⇔ {burn} ⊕ {past} ⊕ {pfv}
i̱š–pasa:–li̱
past–burn–pfv
‘were it burning’

Here, the two inflectional meanings, ‘present’ and ‘irrealis’, are expressed jointly by two morphemes, {past} and {pfv}, that appear individually with these meanings in other, compositional expressions; it is the combination of the two that signifies ‘present irrealis’, just as it is the combination of {shoot} ⊕ {breeze} that expresses ‘chat leisurely’ in (28). Thus, (31) is, like (28), a non-compositional phraseme – in other words, an idiom. The fact that neither of the meanings ‘present’ and ‘irrealis’ corresponds to a component of their morphological expression means that this is a strong idiom.

The expression in (31) demonstrates that morphological idioms can exist at both the levels of derivation and inflection; inflectional idioms differ from derivational idioms, however, in two respects. The first is that inflectional idioms must always be strong idioms, for the simple reason

BDD-A30748 © 2019 Atlantic Provinces Linguistic Association
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-15 01:55:00 UTC)

that if one of a pair (or set) of morphemes, A, expresses the same meaning it has in free expressions, 'A', whereas the other, B, does not, A can be treated as a regular expression of 'A' and abstracted from the representation of the idiom, leaving only B:

(32) 'A' ⊕ 'B' ⇔ {A} ⊕ {C}
⇒ 'B' ⇔ {C}

In this context, C would have to be considered either the regular expression of 'B' (and perhaps homophonous with another C 'C') or it would be a collocational expression of 'B' selected for by A – that is, a compositional morphological phraseme of the type discussed in 3.1.1. Either way, it would no longer be analyzable as an idiom.

The second difference between derivational and inflectional idioms is that inflectional idioms must always be affixal and can never be full. This follows from the simple observation that inflection must always add a discrete and regular meaning to its base, and that once the meaning of the base + inflection becomes non-compositional, the inflection in question has in fact become derivational. Such is the case of the *-s* affix in Spanish *padres* 'parents', which is homophonous with (and historically derived from) the inflectional plural of *padre* 'father'. The two forms, *padres* 'parents' and *padres* 'fathers', coexist in the language, but only in the latter is the *-s* suffix plural inflection. Such would be the fate of any conventionalized combination of a base and an erstwhile inflectional morpheme, eliminating the possibility of having full inflectional idioms.

4. IT'S A SIGN (THAT IT'S A SIGN)

In spite of some easily explicable differences between phrasemes at the phrasal and morphological levels, it seems clear that morphological and phrasal phrasemes have in general the same subtypes and that the basic principles governing phraseologized or conventionalized expressions apply to both. Given what we know about the processes that underlie the diachronic formation and acquisition of fixed expressions, this hardly seems surprising. Phrasal phrasemes are known to result from frequency effects in the distribution of word-level linguistic signs: because of their meanings and common usages, lexical signifiers become conventionally associated with each other, and their context-specific meanings, in the lexicon. If speakers learn such associations for words (lexical signifiers), it is entirely plausible that they could also learn it for morphemes (morphological signifiers). Phraseologization would seem to be a general property of meaning-bearing elements in language, and the hallmark of the linguistic sign.

Interestingly enough, although the notion of routinization and conventionalization of expressions at the phrasal level informs many recent approaches to syntax (e.g., Construction Grammar – Goldberg 1995; Croft 2001), treating affixes as meaning-bearing units is out of step with current fashions in morphology. Word-based approaches to morphology (e.g., Anderson 1992; Haspelmath 2003; Blevins 2006) claim that only whole words are units in the lexicon and that units smaller than words are not representable, the implication being that anything smaller than a word can not be a signifier and that the meaning of affixes can not be recognized out of context – that is, that only words, but not affixes, can be recognized as linguistic signs.

BDD-A30748 © 2019 Atlantic Provinces Linguistic Association
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-15 01:55:00 UTC)

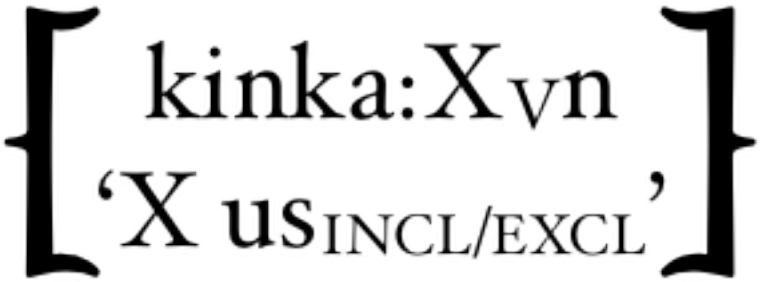
Some of the evidence used in favour of maintaining the integrity of the word come from expressions that in this paper would be analyzed as morphological phrasemes, cases that parallel the Upper Necaxa Totonac sentence in (33):

(33) *kinka:lɑ̨ʔtsiná:n*
kin–ka:–lɑ̨ʔtsin–ya:–n
1obj–pl.obj–see–impf–2obj
i) 'he sees you and me'
ii) 'he sees me and him/them (but not you)'
iii) 'he sees me and him/them and you'

Here we see that one expression, *kin- ka:- -n*, corresponds to three different combinations of persons, depending on whether or not the percept of the seeing event includes, along with the first-person, second and/or third persons. Only the first of the three interpretations, (i), is compatible with the analysis of the word into constituent morphemes, and a word-based morphologist would take this as motivation for treating the expression of the first-person plural object as an abstract schema like that shown in Figure 3:

FIGURE 3

Schematic representation of (33)



Although this representation bears a suspicious resemblance to a linguistic sign (a triplet of signified ('X $us_{incl/excl}$'), signifier (*kinka:- -n*), and syntactics (-X_V-)), it nevertheless avoids explicit parsing of the elements that make up the expression of the first-person plural object, on the grounds that only one of the potential meanings is compositionally expressed by the affixes involved.

This is in some ways very similar to treating the expression in (33) as an inflectional idiom, though the analysis of such expressions as phrasemes has the advantage of not precluding the recognition of the component parts, something which often turns out to be extremely useful. Consider, for example, the sentence in (34):

(34) *kila:lakle:ná:ʍ*
kin–la:–lakle:n–ya:–w

BDD-A30748 © 2019 Atlantic Provinces Linguistic Association
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-15 01:55:00 UTC)

1obj–recip–owe–impf–1pl.sub:incl
i) 'you$_{sg}$ owe us'
ii) 'you guys owe us'
iii) 'you guys owe me'

(34) shows us a non-compositional combination of easily recognizable affixes – *kin-* '1obj', *la:-* 'reciprocal' and *-w* '1pl.sub:incl' – with three interpretations, covering all possible combinations of a second person Actor and a first person Undergoer where one or both is plural. This could be treated schematically along the lines of Figure 3, but doing so obscures important clues to the origins of this conventionalized combination of affixes, which, as noted in Beck (2001), has its roots in extant sociocultural behaviours.

The non-compositional expression in (34) reflects a pattern that can be traced back to proto-Totonacan and seems to be related to a politeness strategy, still in effect, whereby reciprocal forms are used to express a "negative" impact of a second person on a first person. Thus, the expression in (35) is avoided in favour of the expression in (36):

(35) 🔥*kilakle:nɨ́yᵱ*
kin–lakle:n–nɨ–yᵱ
1obj–owe–ben–impf:2sg.sub
'you owe me'

(36) 🌸*la:lakle:ná:ʍ*
la:–lakle:n–ya:–w
recip–owe–impf–1pl.sub:incl
lit. 'we owe each other'

While (36) seems softer and less confrontational than (35), it is also ambiguous, especially when more than two people are involved (i.e., *la:lakle:ná:ʍ* = 'we all owe e.o.', 'I owe you guys and you guys owe me', 'you$_{sg}$ owe us', 'you$_{pl}$ owe me', etc.). Adding *kin-* '1obj' to get the form shown in (34) narrows things down a bit by turning the action back towards the first-person (*kila:lakle:ná:ʍ* 'you owe me', 'you guys owe me', 'you owe us') and, while the expression remains ambiguous, it eliminates the true reciprocal interpretations and limits the endpoint of the event to the first person. The fact that at some earlier point in the development of the language, speakers were able to press-gang the *kin-* prefix in this way seems like a good indication that at some level they were able to parse it and to ascribe to it some meaning associated with first-person deixis – otherwise, adding *kin-* to *la:lakle:ná:ʍ* wouldn't have been an option. What's more, the original use of the highly grammaticized meaning of '1obj' to indicate that the euphemistically reciprocal action was in fact directed towards the speaker(s) seems like a kind of semantic extension (i.e., the signified '1obj' ≠ 'towards/affecting me/us').

Semantic extension of this sort is familiar to us from other studies of the diachronic evolution of words and the creation of morphological forms through the processes of grammaticalization and lexicalization, processes which we have known for a long time don't clearly distinguish phrasal from morphological entities. Grammaticalization, for example, frequently results in full

BDD-A30748 © 2019 Atlantic Provinces Linguistic Association
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-15 01:55:00 UTC)

words evolving over time into affixes (Traugott & Heine 1991; Hopper & Traugott 2003). Indeed, the source of new grammatical elements need not even be genuine words, as in (37):

(37) Eng. *Watergate* > *-gate* 'scandal' (e.g., *Irangate*, *deflategate*, etc.)

In this example, the derivational suffix *-gate* 'scandal' has developed from the second half of a proper name, *Watergate*, which is identifiable as a unit only on phonological grounds but has no meaning or syntactics attached to it. Moving in the opposite direction, Campbell (2001) cites a number of examples of the reverse process in which grammatical, and in some cases morphological, elements become lexical items.

Thus, it seems that a) grammaticalization and lexicalization processes can apply to any type of potential signifier, be that signifier a word, an affix, or a familiar-seeming string, and b) that this association can form linguistic signs of any type – phrasal, morphological, or lexical. The same principles of semantic shift and extension are responsible for turning words into grammatical markers, grammatical markers into words, and abstract grammatical morphemes like *kin-* '1obj' into more "figurative" deictic expressions (and back again). The creation of affixes from words, words from affixes, and combinations of affixes into inflectional idioms appears of be driven by processes that apply to any meaning-bearing elements in language—that is, to linguistic signs.

5. A MORPHOLOGICAL PLATYPUS

In his novel *Lila*, the philosopher Robert Pirsig (1991) discusses what he calls a "platypus," a self-invented problem scientists and others run into when they encounter something in the real world, like a platypus (an egg-laying animal that suckles its young), which contradicts or defies the categorical schemes that they have set up for themselves. As Pirsig observes, the platypus isn't the problem, the problem is the analytical framework that makes the platypus seem problematic. The apparently contradictory nature of affixes like *kin-* that can and can't be parsed strikes me as just that – a platypus based on a caricature of the "classic" morpheme as a segmental "bead on a string" or a node in a tree, each necessarily corresponding to a meaning or a bundle of syntactic or semantic "features." Inflectional idioms like that in (34) are a morphological platypus for this naïve view of the morpheme in that they are segmentable into individual morphs but the morphs do not correspond to the individual meanings they have in other contexts, violating the traditional principle of ascribing a single meaning to a single form.

What the evidence from phraseology has shown us, however, is that the oddness of the parseable form with an unparseable meaning derives only from the failure to recognize that a morpheme is just another type of linguistic sign. Lexical phrasemes – phrases with conventionalized meanings restricted by linguistic or pragmatic context – are widely recognized and well-understood, and the meanings of words in general are known to be flexible and adaptable to syntactic and pragmatic context. No one disputes that analyzable words can participate in non-compositional, non-decomposable expressions, and that this stems from their nature as conventionalized form-meaning associations. Applying the concept of phraseology to morphemes is perhaps novel, but failing to do so leaves us with our morphological platypus and robs us of the simplest and

BDD-A30748 © 2019 Atlantic Provinces Linguistic Association
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-15 01:55:00 UTC)

most obvious explanation for the similarities between phrasal and morphological expressions: both are linguistic signs.

REFERENCES

Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: CUP.

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

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

Bally, Charles. 1951 [1939]. *Traité de stylistique française. Vol. 1*. Paris: Klincksieck.

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

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

Beck, David, & Mel'čuk, Igor A. 2011. Morphological phrasemes and Totonacan verbal morphology. *Linguistics* 49. 175–228.

Blevins, Juliette P. 2006. Word-based morphology. *Journal of Linguistics* 42. 531–573.

Campbell, Lyle. 2001. What's wrong with grammaticalization? *Language Sciences* 23. 113–161.

Croft, William 2001. *Radical Construction Grammar: Syntactic theory in typological perspectiv*e. Oxford: Oxford University Press.

de Saussure, Ferdinand, Baskin, Wade, Meisel, Perry, & Saussy, Haun. 2011. *Course in general linguistics*. New York: Columbia University Press.

Everaert, Martin, van der Linden, Erik-Jan, Schenk, André, & Schreuder, Robert. (eds). 1995. *Idioms: Structural and psychological perspectives*. Hillsdale: Lawrence Erlbaum.

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

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.

Hopper, Paul J., & Traugott, Elizabeth. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.

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

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

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

BDD-A30748 © 2019 Atlantic Provinces Linguistic Association
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-15 01:55:00 UTC)

Mel'čuk, Igor A. 2004. La non-compositionnalité en morphologie linguistique. *Verbum* 26, 439–458.

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

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

Pirsig, Robert M. 1991. *Lila: An inquiry into morals*. New York: Bantam.

Švedova, Natal'ja Ju. (ed). 1980. *Russkaja grammatika, v. I*. Moscow: Nauka.

Traugott, Elizabeth, & Heine, Bernd. (eds.). 1991. *Approaches to grammaticalization: Focus on grammatical markers*. Amsterdam: Benjamins

BDD-A30748 © 2019 Atlantic Provinces Linguistic Association
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-15 01:55:00 UTC)