Igor Mel'čuk

Morphemic and Syntactic Phrasemes

Abstract: A *morphemic phraseme* is a phraseme (= a constrained combination of linguistic signs) composed of morphemes that are part of the same wordform. Like a lexemic phraseme, a morphemic phraseme has a segmental signifier. All logically possible types of morphemic phrasemes are presented and illustrated: morphemic idioms, collocations, nominemes and clichés. Formally, these can be phraseologized complex stems, phraseologized complex affixes and phraseologized wordforms.

A *syntactic phraseme* is a phraseme that includes at least two minimal syntactic subtrees and whose signifier is non-segmental (it involves prosody or an operation). All syntactic phrasemes are idioms. A syntactic idiom must be distinguished from 1) phrases described by means of semantically loaded surface-syntactic relations; 2) phrases consisting of a lexical unit taken together with its actants; 3) lexemic phrasemes consisting of "light-weight" words, such as Rus. ⌜NU I⌝ [X]! lit. 'Well and [X]' = 'What an amazing X!', and 4) lexemic phrasemes with syntactic pecularities. The notion of *fictitious lexeme*, necessary for designating some syntactic idioms (those that are expressed only by prosody), is introduced. An illustrative list of 29 Russian syntactic idioms is presented, as well as the lexical entries for several Russian syntactic idioms.

Keywords: phraseology, morphemic phrasemes, syntactic idioms, fictitious lexemes, Russian syntactic idioms

1 Introduction

The present paper continues my previous work on phrasemes: Mel'čuk 1995, 2012, 2015a, 2015b: 293–362. All relevant notions are introduced in these titles, which the reader is kindly invited to consult for explanations of the necessary terms and formalisms. The definitions of notions that are directly involved in the present discussion are given below.

Let us start with the most general definition of phraseme.

Definition 1: phraseme

> A *phraseme* is a complex linguistic sign $s = s_1 + s_2 + ...$ (= a combination of signs s_1, s_2, ... on the syntagmatic axis) that is *constrained* (= *non-free*): the selection of at least one of the signs s_i by the Speaker depends on other individual signs in the combination.

Igor Mel'čuk, Université de Montréal, Canada, igor.melcuk@umontreal.ca

The best-known phrasemes, or constrained combinations of linguistic signs, are lexemic phrasemes.

Definition 2: lexemic phraseme

> A *lexemic phraseme* is a phraseme consisting of syntactically linked lexemes, that is, a *constrained*, or *non-free*, *phrase*.

Examples: ⌜*under the weather*⌝, ⌜*pull* [N's] *leg*⌝, *pay attention*, *take a shower*, *heavy losses*.

> ☞ The top corners ⌜ ... ⌝ enclose an idiom, see below.

In a lexemic phraseme **s**, all s_i are lexemes, and **s** is a phrase.

Lexemic phrasemes are phrasemes *par excellence*: they are the most numerous and the best studied of all phrasemes. However, languages also use two other major classes of phrasemes: *morphemic phrasemes* and *syntactic phrasemes*, which are the object of the present paper.

Morphemic phrasemes, like lexemic phrasemes, are *segmental signs*: the signifier of a lexemic or a morphemic phraseme is **a string of phonemes** (supplied with a particular prosody). These two classes of phrasemes contrast with syntactic phrasemes, which are non-segmental signs: the signifier of a syntactic phraseme includes—maybe along with a segmental component—autonomous prosody (i.e., prosody unattached to a phonemic string), a configuration of surface-syntactic relations (i.e., non-lexicalized surface-syntactic subtree), grammemes, or an operation. Because of their segmental nature, morphemic phrasemes are more similar to lexemic phrasemes and have been longer known in linguistics than syntactic phrasemes. It seems reasonable, therefore, to begin with morphemic phrasemes.

> **NB** Calling lexemic and morphemic phrasemes segmental signs is *abus de langage*: their components—lexemes and morphemes—are not signs, but sets of signs. Therefore, the term *segmental* is used here as an obvious abbreviation: strictly speaking, lexemic and morphemic phrasemes are, of course, **sets** of segmental signs, but both these types of phraseme are implemented on the linguistic surface as complex segmental signs.

2 Morphemic phrasemes

2.1 The definition of morphemic phraseme

Mel'čuk 1964 proposed a generalization of the notion of lexemic phraseme for it to be applicable to non-free combinations of units of the morphological level, that is, of morphemes. A constrained (= non-free) combination of morphemes within a wordform is a *morphemic phraseme*, or *morphophraseme*.

NB The first explicit mention (known to me) of morphemic phrasemes is found in Pike
1961: 579–581: cf. his terms *affixal idiom* and *affixal collocational complex*. "Word-level
analogs of idioms" are discussed in Dillon 1977: 47 and in Čermák 2007; a detailed char-
acterization of morphophrasemes is given in Mel'čuk 1993–2000: vol. 4, Ch. 9 and Beck
and Mel'čuk 2011.

Definition 3: morphemic phraseme

A *morphemic phraseme* (= *morphophraseme*) is a phraseme consisting of morphemes that
are part of the same wordform.

In a morphemic phraseme **s**, all $\mathbf{s_i}$ are morphemes, and **s** is a wordform or a part
of a wordform (a complex stem or a complex affix).

    Morphemic phrasemes can be considered from a diachronic or from a syn-
chronic viewpoint.

- **From a diachronic perspective**, a morphemic phraseme is:
  - Either a *phraseologized diachronically complex stem*, which can be a *phra-
    seologized diachronically derived stem* (= a phraseologized diachronical
    combination of a stem with derivational affixes) or a *phraseologized dia-
    chronically compound stem* (= a phraseologized diachronical combina-
    tion of a stem with another stem);
  - or a *phraseologized diachronically complex affix* (= a phraseologized com-
    bination of affixes).

In today's language, diachronically complex stems and affixes can remain
complex or become simplexes, depending on the degree of autonomy enjoyed by
their component morphemes: are these morphemes contiguous or not within the
wordform, do they retain their combinatorial properties or not, etc.; see examples
below.

- **From a synchronic perspective**, a morphemic phraseme can only be a
  phraseologized *synchronically complex wordform*, that is, a phraseologized
  combination of a stem with inflectional affixes, where the stem determines
  the selection of the affix expressing the given grammeme(s).

The phraseological properties of morphemic phrasemes are, as indicated above,
similar to those of their lexemic sisters. This should not come as a surprise, since,
as we said, morphemic and lexemic phrasemes are both segmental signs. There-
fore, among morphemic phrasemes, as among lexemic phrasemes, two major
phrasemic subclasses are distinguished:

    semantic-morphemic phrasemes, where the combination of morphemes is
constrained with respect to the corresponding meaning, which is itself freely
constructed by the Speaker; and

conceptual-morphemic phrasemes, where not only the combination of morphemes, but also the underlying meaning is constrained with respect to the corresponding conceptual description of the extralinguistic reality.

In the first class, non-compositional *morphemic idioms* and compositional *morphemic collocations* are distinguished; the second class subdivides into non-compositional *morphemic nominemes* and compositional *morphemic clichés*. All types of morphemic phrasemes can also be constrained pragmatically (= by the situation of their use), that is, be *pragmatemes*. The language-universal typology of morphemic phrasemes is presented in Figure 1.
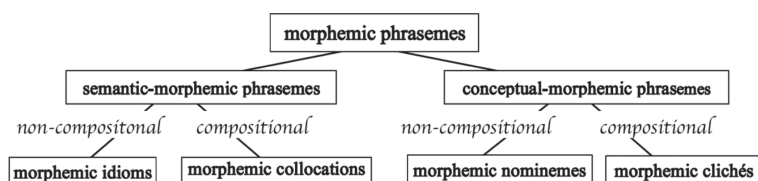

Fig. 1: Universal Typology of Morphemic Phrasemes
(for the universal typology of lexemic phrasemes, see Mel'čuk 2015a: 68)[1].

First, semantic-morphemic phrasemes are presented (Subsection 2.2), then conceptual-morphemic phrasemes (2.3).

2.2 Semantic-morphemic phrasemes

Like semantic-lexemic phrasemes, semantic-morphemic phrasemes come in two major types: morphemic idioms (2.2.1) and morphemic collocations (2.2.2).

2.2.1 Morphemic idioms

Morphemic idioms are found only in diachronic word-formation (= diachronic derivation and diachronic compounding) and diachronic affix-formation, that is, in the historical, at present non-productive creation of new lexemes and new affixes. This is so because synchronic derivation and compounding as well as inflection are by definition too regular to allow for the idiomaticity, i.e., the **semantic** non-compositionality, of produced complex signs.

---

1 Note a change in the terminology: what were called *lexemic phrasemes* in Mel'čuk 2015a and 2020 are now called *semantic-lexemic phrasemes*, and previous semantic-lexemic phrasemes have become *conceptual-lexemic phrasemes*.

> **NB** On the **formal** non-compositionality of complex morphemic signs, see "Morphemic Idioms and Suppletion" below, p. 41.

To put it differently, a morphemic idiom is necessarily stored as a whole (= as a simplex) in the lexicon, if it is a lexeme, or in the grammar, if it is an affix.

Morphemic idioms can be subdivided into the same three subclasses as lexemic idioms: strong idioms, semi-idioms and weak idioms.[2]

> **Terminological remark**
> A morpheme {M} is implemented on the surface by one of its *morphs* **m_i**: thus,
>
> $$\{\text{LEAF}\} \Leftrightarrow \textbf{leaf-}, \textbf{leav-},$$
>
> where the stems **leaf-**(Ø) and **leav-**(*es*) are morphs.
> A *morphemic idiom* ⌐{M_1+M_2}¬ is implemented by a *quasi-morph*; thus,
>
> $$\ulcorner\{\text{FOR}+\text{GET}\}\urcorner \Leftrightarrow \textbf{forget},$$
>
> where the stem **forget-** is—from a diachronic viewpoint—a *quasi-morph*.

The components of a quasi-morph are
– either *morphoids*, if they have semantic links with morphs of the language,
– or *submorphs,* if they do not (Mel'čuk 1993–2000: vol. 4, 249–251).

For instance, in the wordform **conceive**, the elements **con-** and **-ceive** are morphoids; in **forget**, the elements **for-** and **-get** are submorphs.

This analysis is valid only with respect to diachrony. In the modern language, the diachronically complex stems and affixes illustrated below can be simple stems and simple affixes. Thus, in a formal synchronic description of English, {FORGET} and {CONCEIVE} are simple morphemes and **forget-** and **conceive-**, simple morphs.

> ☛ **m** stands for "morph"; within a morphemic idiom, **m** denotes the corresponding *morphoid/submorph*; morphoids/submorphs are separated by pluses: $\textbf{m}_1$+$\textbf{m}_2$; in the examples, lexical stems are printed in CAPS, and affixes, in boldface lowercase characters.

■ Strong morphemic idioms

The meaning of a strong morphemic idiom ⌐{M_1+M_2}¬ does not include the meaning either of {M_1} or of {M_2}: "⌐{M_1+M_2}¬" ⊅ '{M_1}' **and** "⌐{M_1+M_2}¬" ⊅ '{M_2}'.

Strongly idiomatic diachronically complex stems

Strongly idiomatic diachronically derived stems

(1) a. **for**+GET 'X forgets Y' = 'X loses information Y that has been in X's brain': 'forget' ⊅ 'for' **and** 'forget' ⊅ 'get'

---

2 These terms replace, respectively, *full idioms, weak idioms* and *quasi-idioms*, used previously.

    b. WHOPP+**er** 'something extremely large':
      'whopper' ⊅ 'whop [= 'hit hard']' **and** 'whopper' ⊅ '-er'

    Strongly idiomatic diachronically compound stems

(2) a. BLUE+STOCKING 'woman who is too intellectual and not sufficiently feminine':
      'bluestocking' ⊅ 'blue' **and** 'bluestocking' ⊅ 'stocking'
    b. BROW+BEAT 'X browbeats Y' = 'X intimidates Y by stern and/or arrogant behavior':
      'browbeat' ⊅ 'brow' **and** 'browbeat' ⊅ 'beat'

In Modern English, all the stems in (1) and (2) are simplexes. But there also exist strongly idiomatic diachronically derived and diachronically compound stems that remain complex in modern language; for instance, in Dene-Suline (Athabaskan) (Holden 2009):

(3) a. NA +**ne** +HODE+ **s**   +ł +KER 'I beg you' [*'I repeatedly
      ITER 2SG$_{OBJ}$ ask 1SG$_{SUB}$.IMPERF CLASS2 ask   ask you']
      NA-...-HODE-...-KER 'beg' is a strongly idiomatic diachronically derived interrupted (= discontinuous) verbal stem.
    b. K'A +**the** +**Ø** +DA 'S/he lays in ambush'
      arrow 3$_{SUB}$.IMPERF CLASS4 sit
      K'A-...-DA 'lay in ambush' is a strongly idiomatic diachronically compound interrupted verbal stem.

Strongly idiomatic diachronically complex affixes

(4) The German verbal inflectional circumfix **ge**-...-**en** $_{PAST PARTICIPLE}$ (of strong verbs):
    ge+*schrieb*+**en** 'written', **ge**+*flog*+**en** 'flown', **ge**+*schlaf*+**en** '[have] slept'

Diachronically, the circumfix **ge**-...-**en** consists of the prefix **ge-**, which is not productively used in modern German with verbs, and the suffix **-en**, which, outside of this circumfix, marks a verb as either infinitive or 1PL and 3PL.

(5) The Russian nominal derivational circumfix **za**-...-**/j/**(-*e*) 'geographical region behind...'
    This circumfix derives the names of geographical regions situated "behind" a mountain range, a river, or a big lake:[3]
    Za+*kavkaz'*+***j***(-*e*) 'Transcaucasus = region behind (= south of) the Caucasus'
    Za+*volž'*+***j***(-*e*) 'Transvolga   = region behind (= east of) the Volga'
    Za+*bajkal'*+***j***(-*e*) 'Transbaikal = region behind (= east of) Lake Baikal'

---

3 Curiously, the circumfix **za**-...-**j**- is only used to denote a region that is located within Russia; thus, are impossible *****Za***+*rejn'*+***j***(-*e*) 'region behind [= west/east of] the Rhine', *****Za***+*al'p'*+***j***(*e*) 'region behind [= north/south of] the Alps', *****Za***+*and'*+***j***(-*e*) 'region behind [= west/east of] the Andes', etc.

In Modern Russian, the prefix **za-** does not combine with nouns at all; the suffix **-j-** derives depreciative collective nouns, such as BAB′+**j**(-*o*) 'women of the type the Speaker dislikes' (from BABA ≈ 'woman of the type the Speaker dislikes') or DURAČ′+**j**(-*o*) 'fools' (from DURAK 'fool')'.

(6) Three of the Basque (Guipuzcoan) case affixes (Janda and Manandise 1984: 223) are strongly idiomatic diachronically affix complexes:[4]

-rako   DESTINATIVE ['intended for']   : **-ra** ALLATIVE, **-ko** RELATIVE

-rengatik CAUSAL ['because of']   : **-ren** GENITIVE, **-ga** [empty morph], **-tik** ABLATIVE

-rentzat   BENEFACTIVE ['beneficiary for'] : **-ren** GENITIVE, **-tzat** EQUATIVE ['taken for']

---

4 There are two more complex case suffixes in Basque:

-**raino** TERMINAL ['up to'] : -**ra** ALLATIVE, -**ino** TERMINAL and

-**rantz** DIRECTIVE ['toward'] : -**ra** ALLATIVE, -**ntz** DIRECTIVE

They represent, however, a different phenomenon: so-called *parasitic formations*, where an inflectional form is formally built upon another inflectional form. The difference is as follows:

– In a strongly idiomatic complex affix **a** that formally consists of the affixes **a₁** and **a₂**, the signified of **a** is not obtainable as a regular sum of the signifieds of **a₁** and **a₂**: 'a' ≠ 'a₁' ⊕ 'a₂';

   e.g.: Basque -**rako** DESTINATIVE ≠ -**ra** ALLATIVE ⊕ -**ko** RELATIVE

– In a parasitic formation complex affix **b** that formally consists of the affixes **b₁** and **b₂**, the signified of **b** is always equal to the signified of **b₂**: 'b' = 'b₂' (the signified of **b₁** being emptied);

   e.g.: Basque -**raino** TERMINAL = -**ra** ALLATIVE ⊕ -**ino** TERMINAL

A well-known example of a parasitic formation can be found in the case-paradigms of the noun in many Daghestanian languages (Mel'čuk 2006: 457–459). Thus, Archi inflects its nouns for 24 cases, four of which are illustrated below for the noun GEL 'cup':

Case	Number	
	SINGULAR	PLURAL
NOMINATIVE	GEL+Ø+**Ø**	GEL+*um*+**Ø**
ERGATIVE	GEL+Ø+**li**	GEL+*um*+**čaj**
GENITIVE	GEL+Ø+**li**+**n**	GEL+*um*+**če**+**n**
DATIVE	GEL+Ø+**li**+**s**	GEL+*um*+**če**+**s**

NB The alternation -**čaj** ~ -**če** is strictly morphonological: /aj/ → /e/ | __ /C/

Like other languages in the Daghestanian family, a regular Archi noun forms its oblique cases other than the ergative based on the ergative form. Thus, the ergative singular form of GEL is *gelli*, and all the remaining singular forms of the oblique cases are based on the stem *gelli-* (rather than on the radical *gel-*). Likewise, in the plural the oblique case forms are based on the ergative plural, *gelumčaj*, rather than on the nominative plural *gelum*. This is described by the following grammemic rules:

NOM ⇔ {NOM}   GEN ⇔ {ERG} ⊕ {GEN}

ERG ⇔ {ERG}   DAT ⇔ {ERG} ⊕ {DAT}

The genitive, dative, etc. oblique case forms do not express the ergative case, although they contain the ergative suffix (this description is argued for in more detail in Mel'čuk 2008).

■ **Morphemic semi-idioms**

The meaning of a morphemic semi-idiom ⌐{M$_1$+M$_2$}¬ 1) includes the meaning of one of the morphemes {M$_1$} or {M$_2$}, which is not its semantic pivot,[5] and 2) does not include the meaning of the other:

"⌐{M$_1$+M$_2$}¬" ⊃ '{M$_1$}', **and** "⌐{M$_1$+M$_2$}¬" ⊅ '{M$_2$}', **and**
{M$_1$} is not the semantic pivot of ⌐{M$_1$+M$_2$}¬

Semi-idiomatic diachronically complex stems
Semi-idiomatic diachronically derived stems

(7) a. TEAM+**ster** 'professional who drives a truck'
     b. STOPP+**er**   'device designed for plugging an opening'
     c. BIND+**ing**^{1} 'device designed for protecting and keeping together the pages of a book—hard thin plates that...'
     d. BIND+**ing**^{2} 'device designed for fixing a ski boot to the ski'

Semi-idiomatic diachronically compound stems

(8) a. LIGHT+HOUSE 'construction designed for sending signals by light to ships'
     b. SWEET+MEAT   'small food item that is sweet'
     c. PAN+CAKE   'food item that is made from batter cooked in a pan'
     d. LUMBER+JACK 'professional who fells trees for lumber'

Semi-idiomatic synchronic affix complexes (uninterrupted and interrupted)

(9) Kannada (Aronoff and Sridhar 1984):
     IND(icative), NON-PAST, NEG(ation) ⇔ {GERUND} ⊕ {NON-PAST} ⊕ {NEG} ⇔ **uvud⊕-illa**
     MĀḌ+**uvudu** 'doing' ~ MĀḌ+**uvud**+**illa** 'doesn't/won't do'
     IND(icative), PAST, NEG(ation) ⇔ {INF} ⊕ {NEG} ⇔ **-alu⊕-illa**
     MĀḌ+**alu** 'to do' ~ MĀḌ+**al**+**illa** 'didn't do'

---

5 The *semantic pivot of the phrase* L$_1$ + L$_2$ having the meaning 'σ' is the lexeme 1) which has in this phrase its inherent (= non-contextual) meaning 'σ$_1$', 2) such that this 'σ$_1$' is the argument of the remaining part 'σ$_2$' of 'σ'; in other words, 'σ$_2$' = 'σ' – 'σ$_1$' and 'σ$_2$' is a predicate having 'σ$_1$' as its argument, that is, 'σ$_2$'('σ$_1$').

The semantic pivot of a phrase is logically different from the phrase's syntactic head, which, as a rule, expresses the *communicatively dominant component* 'σ̃' of the phrase's meaning 'σ', this component being the minimal paraphrase of 'σ' (Mel'čuk 2001: 29–31). Thus, in the phraseme *take a shower* 'wash oneself under a shower' the semantic pivot is the noun SHOWER, while the syntactic head is the verb TAKE. Note that not every multi-word expression has a semantic pivot. Thus, idioms have no semantic pivot; for instance, the semi-idiom ⌐PRIVATE EYE¬ 'private detective' does not have one.

The *semantic pivot of a morphemic expression* {M$_1$}+{M$_2$} is the same as the semantic pivot of a phrase, but replacing the word *phrase* by the words *morphemic expression* and the word *lexeme* by the word *morpheme*.

(10) Upper Necaxa Totonac (Beck and Mel'čuk 2011: 176)

$1.PL_{SUB}$-EXCL $\Leftrightarrow$ {$1.SG_{SUB}$}, {$1.PL_{SUB}$-INCL} $\Leftrightarrow$ **ik-, -w**

ik+ɬTATÁ+**yā**+**w** 'we.excluding.you sleep' [-*yā*- is a marker of the incompletive aspect]

■ **Weak morphemic idioms**

The meaning of a weak morphemic idiom $\ulcorner\{M_1+M_2\}\urcorner$ 1) includes the meanings of both the morphemes $\{M_1\}$ and $\{M_2\}$, 2) but neither of these is the semantic pivot, and 3) it includes an additional meaning 'A':

$\ulcorner\{M_1+M_2\}\urcorner$ ⊃ '$\{M_1\}$', **and** $\ulcorner\{M_1+M_2\}\urcorner$ ⊃ '$\{M_2\}$', **and** $\ulcorner\{M_1+M_2\}\urcorner$ ⊃ 'A', **and** $\{M_1\}/\{M_2\}$ is not the semantic pivot of $\ulcorner\{M_1+M_2\}\urcorner$.

Weakly idiomatic diachronically complex stems

     Weakly idiomatic diachronically derived stems

(11) a. DIN+**er**   'informal and inexpensive restaurant'

    b. BOMB+**er**   'airplane designed for dropping bombs on targets'

    c. WINE+**ery**   'business where wine is produced and stored'

    d. Rus. SPASA+**tel′** lit. 'sav+er' = 'lifeguard; first responder' =   'professional who saves people in extreme situations'

    e. Ger. LEICHEN+BESCHAU+**er** lit. 'corpse examiner' =   'professional who examines corpses in order to make out death certificates'

    f. Ger. BRIEF+TRÄG+**er** lit. 'letter carrier' = 'professional who delivers mail'

NB Examples (11e–f) present stems derived from nominal compounds.

     Weakly idiomatic diachronically compound stems

(12) a. FEED+BACK   'information about the results of an action [by X on Y] [that is] fed back [to X]'

    b. SNOW+SHOES   'device [designed for X] to walk on snow—flat frames to be attached under X's shoes'

Morphemic idioms and suppletion (Mel'čuk 1993–2000: vol. 4, 403)

A morphemic idiom is defined by its **semantic** non-compositionality—that is, its signified is not a regular (i.e., compositional) union of the signifieds of its component morphemes. However, formally it is quite regular, that is, compositional. At the same time, there are complex morphological signs of, so to speak, an inverse nature: they are semantically compositional, but formally non-compositional. Thus, the English wordform **has** is semantically compositional, since its signified is regularly constructed out of the signifieds of its virtual components:

$$\text{'has'} = \text{'have'} \oplus \text{'IND.PRES'} \oplus \text{'3'} \oplus \text{'SG'}.$$

But its signifier is not regularly constructed: it "should" be ***haves**, but it is **has**. This is a well-known linguistic phenomenon: the wordform **has** is a *strong mega-morph suppletive* with respect to the stem **have**-. I am not in a position to develop this point any further here (on suppletion, see Mel'čuk 2006: 405–467), and I will limit myself to emphasizing the following fact:

> Either the signified of a phraseologized morphological complex sign can be a non-regular (= compositional) combination of the signifieds of its components, or its signifier can be a non-regular combination of the signifiers of its components.

The first case gives us morphemic idioms, and the second, pairs of suppletive units—a strong megamorph and the corresponding stem. In one of his last papers, Weinreich (1969: 43) defines an idiom as a phraseme whose meaning is "suppletive [sic!—IM] with respect to the sum of the meanings of its components," i.e., Weinreich treats phraseologization as suppletion in the domain of meaning. Inversely, we can say that suppletion is phraseologization in the domain of form.

　　The properties of morphemic idioms and those of strong megamorphs can be presented in parallel:

A **morphemic idiom** $\ulcorner\{M_1+M_2\}\urcorner$ is implemented by a non-elementary segmental sign **m1+m2** whose signifier is regularly representable in terms of the signifiers of its components, but whose signified is not regularly representable in terms of the signifieds of its components.	A **strong megamorph m̃** = **m1.m2** is a non-elementary segmental sign whose signified is regularly representable in terms of the signifieds of its components, but whose signifier is not regularly representable in terms of the signifiers of its components.
Schematically: 'm1+m2' ≠ 'm1' ⊕ 'm2' /m1+m2/ = /m1/ ⊕ /m2/ **Example** 'for+get' ≠ 'for' ⊕ 'get' /fɔr+get/ = /fɔr/ ⊕ /get/	Schematically: 'm̃' = 'm1' ⊕ 'm2' /m̃/ ≠ /m1/ ⊕ /m2/ **Example** 'is' = 'be' ⊕ '{IND.PRES}' ⊕ '{3SG}' /ɪz/ ≠ /bī/ ⊕ /Ø/ ⊕ /z/

These close ties between suppletion and phraseologization are theoretically important.

2.2.2 Morphemic collocations

☞ The collocation base is printed in SMALL CAPS, and the *collocate* (in this case, the derivational affix or dependent compound component), in boldface lowercase **characters**.

Morphemic collocations are found in diachronic and synchronic derivation and compounding, as well as in inflection.

Collocational complex stems

      Collocational diachronically derived stems

Here are two stock examples of diachronic derivational collocations.

(13) Names of inhabitants

    LONDON+**er**, BOSTON+**ian**, MUSCOV+**ite**, VIENN+**ese**, DAMASC+**ene**, HYDERABAD+**i**, SYDNEY+**sider**

One of the suffixal morphemes {PERSON WHO LIVES IN L} is selected by the Speaker as a function of the stem, which is the name of a city or a town.

(14) Action nouns

    ACCEPT+**ance**, ACKNOWLEDG+**ment**, OBSERV+**ation**, INTRUS+**sion**, REFUS+**al**, OUST+**er**, WEIGHT+**ing**

Now, three more complex examples of diachronically derivational collocations.

(15) Russian

    a. PAS+**tux** lit. 'pastur+er' = 'shepherd'
        The noun is diachronically derived from PAS(-*ti*) 'pasture$_{(V)}$ [trans.]' by the agentive unisuffix[6] **-tux** 'person who...'.

---

6 A unisuffix, which is encountered just with one stem, is a suffixal analog of a unilexeme. Unilexemes are of two kinds:

—A unilexeme is part of an idiom and used only in this idiom; it is semantically empty. These unilexemes can be called idiom unilexemes. For instance:

KITH,   in ⌜KITH AND KIN⌝ 'close friends and relatives'
RUNCIBLE,   in ⌜RUNCIBLE SPOON⌝ 'three-pronged fork, curved like a spoon and having a cutting edge'
SPIC and SPAN, in ⌜SPIC AND SPAN⌝ 'clean and bright, like brand-new'
FRO,   in ⌜TO AND FRO⌝ 'in one direction and then back again'

—A unilexeme is part of a collocation and cannot be used without its base or its collocate; it is semantically full. These unilexemes can be called collocational unilexemes. For instance:

AQUILINE,   used only with NOSE

    b. POP+**ad'j**(-*a*) lit. 'Orthodox.priest+wife'
       -ad'j- is also a unisuffix.
    c. TOPOR+**išč**(-*e*) lit. 'axe+handle'

Russian has two homonymous derivational suffixes **-išč-**:
– the semi-productive augmentative suffix **-išč[1]-** 'extremely big', as in NOS+**išč**(-*e*) 'extremely big nose', KOMNAT+**išč**(-*a*) 'extremely big room', etc.;
– the unproductive, but frequent locative suffix **-išč[2]-** 'place where …', as in POŽAR 'fire'+**išč**(-*e*) 'place where there has been a fire', STREL′B 'shoot'+**išč**(-*e*) 'shooting range', etc.

The suffix **-išč-** carries the meaning 'handle' only in the lexeme TOPORIŠČE.
(KNUT+**ovišč**(-*e*) 'whip handle' contains a different, albeit similar, suffix: -**ovišč-**.)

      Collocational diachronically compound stems

(16) German
    a. **Haus**+TÜR lit. 'house door' = 'the door that is the main entrance to the building' = 'front door'
    b. **Schmerz**+**ens**+GELD lit. 'pain money' = 'money paid as compensation for personal injuries'

Collocational synchronically complex wordforms (inflection)

In a language with rich non-agglutinative morphology, where nouns of different types have formally different declensions and verbs of different types have different conjugations, each declined or conjugated wordform (that is, each inflectional form) is a morphemic collocation: its stem is the base and the inflectional suffix is the collocate, since it is selected for the given grammemic combination as a function of this stem. Four examples follow.

(17) Russian declensions
    I[st]   — SG, NOM KNIG+**a** 'book'
    IInd MASC — SG, NOM ROG+**Ø** 'horn'; NEU — SG, NOM BOLOT+**o** 'swamp'
    IIIrd   — SG, NOM NOČ′+**Ø** 'night'

PYRRIC,   used only with VICTORY
HEADWAY,   used only with MAKE
STRINGS,   used only with PULL
All unisuffixes are collocational unisuffixes.

(18) Spanish conjugations

     Ist — INF NAD+**a**(-*r*) '[to] swim'
    IInd — INF COC+**e**(-*r*) '[to] cook'
  IIIrd — INF MOR+**i**(-*r*) '[to] die'

(19) Russian verbal aspect prefixes (six out of 15 aspect prefixes are given)

IMPERFECTIVE	PERFECTIVE			IMPERFECTIVE	PERFECTIVE		
ČITAT′	~	**pro**+ČITAT′	'read'	STROIT′	~	**po**+STROIT′	'build'
BUDIT′	~	**raz**+BUDIT′	'wake'	PIT′	~	**vy**+PIT′	'drink'
DELAT′	~	**s**+DELAT′	'do'	POIT′	~	**na**+POIT′	'make drink'

The perfective forms are compositional: all the prefixes express only 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. This means that each verb has to be marked in the lexicon for the particular perfective prefix morpheme it takes. The choice of the perfective prefix for each radical is thus severely constrained, but the resulting form is compositional. This is characteristic of collocations: the stem is the base of the collocation, as well as its semantic pivot.

(20) Yasin-Burushaski noun plural suffixes (Berger 1974: 15–20; twenty out of about 70 plural suffixes are given)

SINGULAR		PLURAL			SINGULAR		PLURAL		
THÁM	~	THÁM	+**u**	'king'	HÍR	~	HUR	+**í**	'man'
PÁQU	~	PÁQU	+**mu**	'bread'	DÍU	~	DIW	+**ánc**	'demon'
AIŽDAHÁR	~	AIŽDAHÁ	+**išu**	'dragon'	ASQÓR	~	ASQÓR	+**iŋ**	'flower'
TÁΓ	~	TAΓ	+**ášku**	'branch'	HÁRČ	~	HARČ	+**óŋ**	'plow'
TÁL	~	TÁL	+**ǯu**	'pigeon'	TÍŠ	~	TIŠ	+**míŋ**	'wind'
DÁN	~	DAN	+**ǯó**	'stone'	WAZÍIR	~	WAZÍIR	+**tiŋ**	'minister'
DUŠMÁN	~	DUŠMÁ	+**yu**	'enemy'	GÚS	~	GUŠ	+**íŋa**	'woman'
ČÁR	~	ČAR	+**kó**	'rock$_{(N)}$'	GÓṬ	~	GOṬ	+**ó**	'[a] mute'
HÚK	~	HUK	+**á**, +**ái**	'dog'	ḌÍM	~	ḌÍM	+**a**	'body'
ÚRK	~	URK	+**á**, +**ás**	'wolf'	TÚR	~	TUR	+**iáŋ**	'horn'

The plural forms of Burushaski nouns are semantically compositional, but formally unpredictable: for each individual radical (which is the base of the morphemic collocation and its semantic pivot), the corresponding plural suffix has to be specified in the lexicon. The distribution of the plural morphemes 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.

2.3 Conceptual-morphemic phrasemes

2.3.1 Morphemic nominemes

Morphemic nominemes are the proper names of individual places, events, human groups, etc., and, therefore, like lexemic nominemes, they represent little interest for a language lexicon: their place is rather in an encyclopedia.

Diachronically derived morphemic nominemes

(21) Russian
   a. ALEKSANDR+**ov**(-Ø) lit. 'Alexander's': a city in Russia, named after some Aleksandr.
   b. EŽOV+**ščin**(-*a*) lit. 'Ezhov era': the period in the USSR (1936–1939) named after the then Minister of State Security Ežov, when Stalinist terror reached its peak.
   c. STAXANOV+**c**(-*y*) lit. 'Stakhanov+ite(s)': members of the mass movement of workers in the USSR in the 1930s aimed at working harder and producing more; the movement was named after its initiator, a miner called Staxanov.

Diachronically compound morphemic nominemes

(22) a. German
   NEU+STADT lit. 'new city': a city in Germany (cf. NOV+GOROD lit. 'new city': a city in Russia).
   SCHWAN+GAU lit. 'swan district': a district in Germany.
   b. English
   GREEN+LAND; BURN+SIDE [a city in the US]; NEW+TOWN [a city in the US]

2.3.2 Morphemic clichés

For the time being, I have examples for three subclasses of morphemic clichés: morphemic nicknames, morphemic termemes and morphemic formulemes (but no examples for morphemic sentencemes[7]).

---

7 A morphemic sentenceme should be a morphemic cliché that has a generic abstract referent, i.e. it denotes a class of situation; examples of lexemic sentencemes are as follows: *A drowning man will clutch at a straw* and *If you play with fire, you'll get burned*. A morphemic sentenceme is theoretically possible only in a polysynthetic language, where one wordform can correspond to a full sentence.

■ Morphemic nicknames

A morphemic *nickname* is a morphemic cliché that has a specific concrete referent, that is, it refers to an individual (in the logical sense). But unlike a nomineme, a nickname has a meaning: it not only identifies its referent, but tells us something about it; its lexical components are semantically full (= meaningful) morphemes. Nevertheless, morphemic nicknames, just like their lexemic counterparts, belong in an encyclopedia.

(23) German
- a. KRISTALL+NACHT lit. 'crystal night': the night of Nov. 9–10 1938 in Germany when massive pogroms against Jews were organized by the government and when mobs broke Jewish shop windows.
- b. BRAUN+HEMD(-*en*) lit. 'Brown Shirts': members of the paramilitary wing of the National-Socialist party in the Nazi Germany, who wore brown shirts.

■ Morphemic termemes

A morphemic *termeme* is a morphemic cliché that has a generic concrete referent, that is, it denotes a particular class of individuals (in the logical sense). This means that a termeme corresponds to a technical term.

(24) German
- a. AUSLANDS+MINISTERIUM lit. 'abroad ministry' = 'ministry of foreign affairs'
- b. TEE+ROSE 'tea rose'
- c. BLIND+DARM lit. 'blind gut' = 'caecum'
- d. ROT+WEIN 'red wine'
- e. WEISS+WURST lit. 'white sausage' = 'veal sausage'

■ Morphemic formulemes

A formuleme is a cliché that has a specific abstract referent, that is, denotes a particular situation.

(25) a. THANK+**s**! and Serb. HVAL+**a**! 'Thanks!'
- b. Ger. ENTSCHULDIG+**ung**! lit. 'excusing' = 'Sorry!'

Many morphemic formulemes are constrained pragmatically, that is, by the situation of their use. Such formulemes are, at the same time, pragmatemes (Mel'čuk 2020):

- c. STOP+Ø [**on a traffic sign**] (= [to] STOP$_{IMPER}$)
- d. Fr. TIR+**ez** 'Pull' [**sign on a door in a public building**] (= TIRER$_{IMPER, 2, PL}$ 'pull').
- e. Rus. LOŽ+**is'**! lit. 'Lie.down!' = 'Duck!/Take cover!' [**in a situation of shooting**]

(*Ložis'* = LOŽIT'SJA_{IMPERF, IMPER, 2, SG} 'lie down')[8]

and

Pol. PADN+**ij**! lit. 'Drop down!' = 'Duck!/Take cover!' [**in a situation of shooting**]

(*Padnij* = PAŚĆ_{PERF, IMPER, 2, SG} 'drop down')

    f.   Rus. PRIVET+**Ø**! lit. 'greeting_(N)' [**in a situation of greeting**] = 'Hi!'

    g.   Upper Necaxa Totonac (D. Beck, personal communication)

WILÁ+**ya**? lit. 'Are you sitting?' [**in a situation of greeting somebody sitting**] = 'Hi!'

PIN+**pá**?   lit. 'Are you going?' [**in a situation of greeting somebody walking**] = 'Hi!'

KUWÍN+**i**! lit. 'Be.late.morning 2SG_{SUB}.PERF' [**in a situation of greeting somebody in the morning**] = 'Good morning!'

2.4 The place of morphemic phrasemes in language

Interestingly, the following statement seems to hold:

> Many lexemes of a language are in fact diachronic morphemic phrasemes.

How many? Strictly speaking, nobody knows, but we can try a rough estimate.

Among the thousand words of basic English roughly 10% are diachronic morphemic idioms. Here is a sample count performed for five letters (A, C, H, P and S):

A: 14 out of 48		C: 8 out of 76		H: 3 out of 49	P: 5 out of 78	S: 10 out of 136	
across	always	careful	cloudy	healthy	photograph	someone	substance
actor	another	careless	comfortable	hers	pleased	something	successful
active	anyone	central	computer	holiday	probably	sometimes	suitable
activity	anything	clothes	cupboard		produce	student	sunny
alone	anytime				provide	subject	support
along	around						
already	away						

Total: 40 words out of 387 are diachronically derived, which makes > 10%.

In the general vocabulary, the proportion of diachronically complex (= diachronically derived or compound) lexemes must be—and is—much higher. Thus, if we arbitrarily take five pages from the *Longman Dictionary of Contemporary English* (1978), the picture is as follows:

p. 150: 29 words ~ 10 diachronically complex (= derived or compound)

p. 250: 35 words ~ 23 diachronically complex

---

8 Note that the form *Ložis'!* is in the singular, even when addressed to many people; the same is true about the Polish verb form *Padnij!*

p. 350: 34 words ~ 20 diachronically complex
p. 460: 22 words ~ 15 diachronically complex
p. 551: 37 words ~ 11 diachronically complex

Total: 79 out of 157 words are diachronically derived or compound; this is already ≈ 50%! Our rough calculations suggest that almost half of the lexical stock of a language are diachronically complex lexemes. Thus, morphemic phrasemes have an important place in the vocabulary of a language.

It is true that for a synchronic description of language **L** morphemic phrasemes are, logically speaking, irrelevant: derived and compound stems are presented in **L**'s lexicon as simplexes. Nevertheless, from a pedagogical viewpoint, the indication of diachronic derivation and compounding for a synchronic simplex seems quite useful: for instance, it gives a human user additional knowledge about the semantic range of the lexeme he considers.

Having discussed morphemic phrasemes, I can now move to syntactic phrasemes.

3 Syntactic phrasemes

3.1 Introductory remarks

As far as I know, the notion of *syntactic phraseme* as one of the three major classes of phrasemes was formally introduced in Mel'čuk (1987: 645). Syntactic phrasemes are in contrast to both *lexemic* and *morphemic phrasemes*. As a typical example of an English syntactic phraseme (more precisely, a *syntactic idiom*; for the definitions, see Subsection 3.3.1 below), the following tautology-like expression (which is by no means a tautology!) can be cited (Wierzbicka 1987):

(26) [Xs] ⌜ARE L(X)s⌝ ≈ 'Humans X, in conformity with their nature, do some undesirable things that are not so bad and can be put up with'.

☛ Reminder: a variable in square brackets stands for a semantic actant slot; L(X)—a bound lexemic variable—denotes the lexeme L that expresses X; the top corners ⌜ ... ⌝ enclose an idiom.

This syntactic idiom is implementable as *Boys$_X$ are boys$_{L(X)}$; Wives$_X$ are wives$_{L(X)}$; Poets$_X$ are poets$_{L(X)}$*; etc. It expresses, in Wierzbicka's terms, "tolerance for human qualities." Wierzbicka (1987) offers detailed semantic descriptions of many "tautologically organized" syntactic idioms in several languages; see also Rhodes (2009).

The deep-syntactic [DSynt-] and surface-syntactic [SSynt-]structures of this syntactic idiom are as follows:

DSyntS: $[X_{PL}] \leftarrow I - \ulcorner BE\ L(X)_{PL} \urcorner$ ⇔

SSyntS: $[X_{PL}] \leftarrow$**subjectival**$-BE_{IND, PRES, NON-PERF, NON-PROGR}-$**copular-completive**$\rightarrow L(X)_{PL}$

This idiom is not a segmental sign: its signifier contains, along with the copula BE, an **operation**—the duplication of the actant X; this operation is represented by the bound lexemic variable L(X). Since such non-segmental signs are not very well known in the literature, it is useful to give some more examples. Thus, here is a Russian syntactic phraseme (also a syntactic idiom):

(27) $\ulcorner$BYT'**I.3** VSEM $L(X)_{(N)PL, DAT}\urcorner$ $[X_{(N)\mathbf{n}, NOM}]$ 'This is the most outstanding X of all Xs'[9]
    Askrej i Arsija na Marse – vsem goram$_{L(X)}$ *gory*$_X$
    'Ascraeus Mons and Arsia Mons on Mars are the most outstanding mountains$_X$ of all mountains$_{L(X)}$'.

    ☞ The subscript **n** stands for "number"; it refers to the number of the subject of the copula BYT'**I.3**.

This idiom's DSynt- and SSynt-structures are here:

DSyntS: $\ulcorner$BYT'**I.3** VSEM $L(X)_{(N)PL, DAT}\urcorner - II \rightarrow [X_{(N)\mathbf{n}, NOM}]$ ⇔

SSyntS:
    $\overbrace{\qquad\qquad}^{\textbf{attributive}}$
BYT'**I.3** 'be'—**copular-completive** $\rightarrow [X_{(N)\mathbf{n}, NOM}]$ VSE 'all'$\leftarrow$**modificative**$-L(X)_{(N)PL, DAT}\urcorner$

Now come three more Russian syntactic idioms. These expressions are of a very different nature, and my choice is deliberate: they demonstrate the astonishing structural variety of syntactic idioms.

(28) a. $[X \ulcorner$XOT'$\urcorner L(INTENS(X))_{(\tilde{v})IMPERF, IMPER, 2, SG}]\urcorner$[10]
    {*Mne stalo*} [*stydno*$_X$]**, *xot' von begi*$_{L(INTENS_X)}$** ⟨*xot' v pogreb prjač'sja*; ...⟩
    '{I was} ashamed$_x$, even ready to run away ⟨to hide in a cellar; ...⟩.

---

9 For different lexemes of the verb BYT' 'be' in Russian, see Mel'čuk (2019: 2.2, 11–16). BYT'**I.3** means 'be an element of class...; BYT'**I.2** in (28b) means 'be identical to'.
10 Three linguistic comments:

    1. XOT' ≈ 'at least' is an adverb that cannot be translated directly into English. In these phrasemes, XOT' means something like 'even ready to ...'.

    2. "$L(INTENS(X))_{(\tilde{v})IMPERF, IMPER, 2, SG}$" stands for a verb phrase whose head is a verb in the 2nd person singular of the imperfective imperative, while its lexical filling may be anything provided that the resulting expression is semantically fit to serve as an intensifier of X. An $L(INTENS(X))_{(\tilde{v})IMPERF, IMPER, 2, SG}$ phrase can be considered to be a peculiar value of the lexical function Magn applied to X.

b. $[X_{NOM}]$ ⌐ÈTO BYT′**I.2**$_{PRES, 3, SG}$ L(X)$_{NOM}$⌐

 [*Vojna*$_X$] – ***èto vojna***$_{L(X)}$ lit. 'War$_X$ it [is] war $_{L(X)}$'. ≈ 'War is war'.

 [*Matematika*$_X$] – ***èto matematika***$_{L(X)}$ 'Mathematics is mathematics'.

 [*Fakty*$_X$] – ***èto fakty***$_{L(X)}$ 'Facts are facts'.

c. «WILL.BE.PUNISHED»

 [*On*$_X$] [*u menja*$_Z$] [*budet valjat′sja*$_Y$ *na divane*]*!*

 lit. 'He $_X$ at me$_Z$ will lie.around$_Y$ on the.sofa!' =

 'If he continues to lie around on the sofa he will be severely punished by me!'

NB In (28a) and (28b) the surface "manifestations" of the syntactic idioms are boldfaced. But in (28c) nothing can be boldfaced, because the signifier of this syntactic idiom does not contain segmental components: the meaning of threat is expressed exclusively through prosody attached to this particular surface-syntactic relation configuration. This idiom has to be represented (in the deep-syntactic structure) by a fictitious lexeme: «WILL.BE.PUNISHED» (Subsection 3.2 below, (29)).

Expressions of the type presented in (28) have been thoroughly described: for instance, Šmelëv (1960), Švedova (1960: 269-279), Švedova, ed. (1970: 558, 563–565; 1980: 85, 385–386), Wierzbicka (1987), Kajgorodova (1999), Lim (2001), Iomdin (2006a, 2006b, 2010, 2013, 2017), Kopotev (2008), Kopotev and Steksova (2016), Vilinbaxova and Kopotev (2017), Avgustinova and Iomdin (2019), Dobrovol′skij *et al.* (2019). These expressions have been given different names: "(bound) constructions," "phraseoschemata," "syntactic phraseologisms," or, as a cover term, "syntactic phrasemes," or else "syntactic idioms." However, the expression *syntactic phrasemes/idioms* seems to be abused: it is often interpreted too widely and too vaguely, namely as referring to any multiword expression that features syntactic peculiarities. In a similar vein, Construction Grammar (Goldberg 1995; Goldberg and Jackendoff 2004; Raxilina ed. 2010) considers as *constructions* not only phrasemes, but also phrases manifesting government of syntactic actants, that is, phrases of the form L$_1$–**synt**→L$_2$, where the lexeme L$_2$ is an actant of the lexeme L$_1$. As a result, linguistic phenomena of quite a different

───────

3. Russian has a number of lexemic idioms with XOT′:

 ⌐XOT′ UBEJ(TE)⌐ 'You can kill me, but I am unable to do this'

 ⌐XOT′ PLAČ′⌐ 'You can cry, but this won't help'

 ⌐XOT′ IZ PUŠEK STRELJAJ⌐ 'You can fire big guns, but this won't wake this person'

 [Y$_{DAT}$] ⌐XOT′ KOL NA GOLOVE TEŠI⌐ 'You can whittle off a pole on Y's head, but Y won't understand', etc.

These lexemic idioms must, of course, be distinguished from the syntactic idioms considered in this paper. Cf. also note 19, p. 69.

nature get mixed up, and this leads to incorrect and/or unnatural descriptions. To correct this drawback the present paper proposes and illustrates a rigorous definition of the notion *syntactic phraseme* (see Section 3.3.1).

3.2 Typical Russian syntactic phrasemes (= syntactic idioms)

Consider three Russian sentences (the subscripts $_X$, $_Y$, and $_Z$ designate the semantic actants of the expressions under analysis):

(29) a.   – *Ty$_X$ poguljaeš'$_Y$ u Ivana$_Z$ po nočam!*
     lit. 'You$_X$ will.go.for.a.walk$_Y$ at Ivan$_Z$ at night!' [**threat**] =
     'I signal: if you go for a walk at night, you will be severely punished by Ivan'.

    b.   – *Kak že, ty$_X$ poguljaeš'$_Y$ u Ivana$_Z$ po nočam!*
     lit. 'Well, well, you$_X$ will.go.for.a.walk$_Y$ at Ivan$_Z$ at night!' [**sarcastic negation**] =
     'I signal: it is impossible that you will be allowed by Ivan to go for a walk at night'.

    c.   – *Ne somnevajsja, ty$_X$ poguljaeš'$_Y$ u Ivana$_Z$ po nočam!*
     lit. 'Don't doubt, you$_X$ will.go.for.a.walk$_Y$ at Ivan$_Z$ at night!' [**assurance**] =
     'I signal: no doubt that you will be made by Ivan to go for a walk at night'.

All three sentences are *signalatives* (Mel'čuk 2001: 242–251, 354–256). That is, they do not communicate statements about facts or entities in the extralinguistic world, that is, statements that can be negated or interrogated. A signalative *signals* a mental state of the Speaker ('I signal that...') and represents his particular speech act: (29a) is the Speaker's threat that is intended to prevent the action 'you go for a walk at night'; in (29b), the Speaker sarcastically negates the possibility of this action; on the contrary, (29c) gives the Speaker's assurance of its future realization.[11] The three sentences in (29) consist of the same lexemes, linked by the same syntactic relations; they differ only in prosody. That is, the meanings that distinguish these sentences are expressed exclusively prosodically; the prosody of each one of them is invariably associated to a particular predicate meaning '$\sigma_{(29a)}$'/'$\sigma_{(29b)}$'/'$\sigma_{(29c)}$'. The lexemes with subscripts X, Y and Z in (29) express the *semantic actants* of these meanings, while each meaning '$\sigma_{(29a)}$'/'$\sigma_{(29b)}$'/'$\sigma_{(29c)}$' itself is carried by the prosody of the sentence—that is, it is

---

11 The signalative character of these sentences is manifested, in particular, in that they (taken with the corresponding prosody) cannot be syntactically subordinated to a verb via the complementizer что 'that'.

expressed by a non-segmental means. Thus, sentences (29) contain linguistic signs of quite a particular type. A question arises naturally:

> How should the signs carrying the meanings '$\sigma_{(29a)}$'/'$\sigma_{(29b)}$'/'$\sigma_{(29c)}$' in sentences (29) be described formally?

Let us start with the type of sign in sentence (29a).
– Its *signified* is ordinary, the same as with most linguistic signs: the semantic representation [SemR] of a meaning.
– Its *signifier*, on the contrary, is quite special. On the morphological level it is not a phonemic string: the lexemic variables X, Y and Z, linked by particular surface-syntactic relations [SSyntRels], stand for *the actant slots* of the meaning '$\sigma_{(29a)}$' ≈ 'Z will punish X for Y(X)', not in any way for the meaning '$\sigma_{(29a)}$' itself. This meaning is expressed only by a particular prosody. Such a signifier must be presented (in the lexicon) in the form of a prosodic structure.
– Its *syntactics* is also special. On the one hand, it specifies the configuration of surface-syntactic relations (= a subtree) on which the prosodic signifier is to be imposed. On the other hand, it indicates that the resulting expression is a full sentence and a signalative; as for its style, it is colloquial.

Let us call the sign under consideration «WILL.BE.PUNISHED». This expression «WILL.BE.PUNISHED» is a *fictitious lexeme* (see Section 3.3.3).

Here is a formal representation of this sign.

«WILL.BE.PUNISHED»
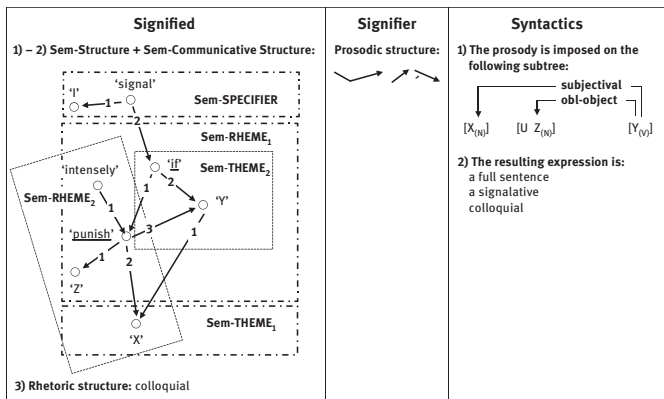
☞ In the prosodic structure of this sign's signifier, the arrows show the movement of the intonation, and the symbol ′ indicates a particular type of strong accent.

Fig. 2: Formal representation of the Russian non-segmental sign «WILL.BE.PUNISHED»

Now, what is the class of linguistic signs to which the sign «WILL. BE.PUNISHED» belongs? The answer depends on this sign's signifier, since the signifieds of all types of linguistic signs are of the same logical and formal nature. The signifier of the «WILL.BE.PUNISHED» sign is a particular prosodic structure to be imposed on the string of lexemes in order to form a sentence; but these lexemes are not part of the sign: some of them are the sign's actants, and the others are dependents of these actants. The meaning carried by said prosodic structure, that is, the signified of the sign «WILL.BE.PUNISHED», shown as the SemR in Figure 2 (≈ 'I signal that if X does Y, X will be severely punished by Z for Y'), is not built in a regular way out of the signifieds of its components. To put it in other words, the meaning of this sign cannot be regularly distributed among the components of its signifier, that is, between chunks of its prosody. Therefore, the sign «WILL.BE.PUNISHED» is not compositional and thus it is an idiom. And since its signifier is not segmental, it is a syntactic idiom. This idiom is, as I have said, denoted with a fictitious lexeme; and a fictitious lexeme, like any genuine lexeme, has its own lexical entry, which is given below.

> **NB** The lexical entry of a lexemic idiom features two additional zones with respect to the lexical entry of a single lexeme:
> – The zone of **surface-syntactic implementation**, which specifies the SSynt-tree of the idiom.
> – The zone of **deep-morphological implementation**, which contains the necessary data on the linear order of the idiom's lexemic components, on possible gaps between these components, etc.
>
> The lexical entry of a syntactic idiom also has these two zones: they likewise specify the SSynt-tree of the idiom and the linear ordering of its lexemic components; however, in the case of a syntactic idiom whose signifier consists only of prosody they specify the SSynt-structure and the ordering not of the idiom components themselves, but of the components of the resulting expression—the idiom plus its actants. In the illustrative lexical entries given below, the deep-morphological implementation zone is not shown.

«WILL.BE.PUNISHED», fictitious lexeme that represents a syntactic idiom; with its actants, it forms a full sentence, which is signalative; **colloquial**

Definition

'[X] «WILL.BE.PUNISHED» [for Y by Z]' = 'I signal that if X does Y, X will be severely punished by Z for doing Y'

Government Pattern

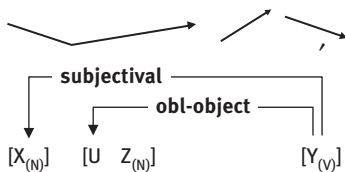| X ⇔ **I**	Y ⇔ **II**	Z ⇔ **III**
1. N_{NOM}	1. $V_{IMPERF, FUT}$ 2. *po*+$V_{PERF, FUT}$	1. *u* 'at' + N_{GEN}

Ivan$_X$ u novogo direktora$_Z$ budet$_Y$ p'janstvovat'! ⟨= *popjanstvuet!*⟩
lit. 'Ivan$_X$ at new director$_Z$ will$_Y$ be.drinking!' =
'If Ivan drinks, the new director will severely punish him'.

Surface-Syntactic Implementation

Surface-Syntactic Implementation

[Diagram: curved lines above, with labels]
subjectival
obl-object
$[X_{(N)}]$ $[U$ $Z_{(N)}]$ $[Y_{(V)}]$

Once again, $[X_{(N)}]$, $[Y_{(V)}]$ and $[Z_{(N)}]$ are the semantic actants of the idiom «WILL.
BE.PUNISHED», not its "variable lexical parts," as they are often characterized.

The signs of the «WILL.BE.PUNISHED» form, where the signifier is just the prosody, are by no means the only variety of syntactic idioms. Another (and quantitavely overwhelming) type of Russian syntactic idiom is presented in (30); these idioms are non-segmental because of a bound lexemic variable in their signifier:

(30) a. {*Na pervyj vzgljad,*} [*derevo*$_X$] *bylo **kak derevo***$_{L(X)}$
       lit. '{At first sight,} [tree$_X$] was as tree$_{L(X)}$'. =
       'At first sight, this tree was **quite an ordinary** tree'.
   b. [*Nam*$_Y$] [*obed*$_X$] **Ø$^{BYT'I.3}$ 'be'** *ne v obed*$_{L(X)}$ {*kol' xozjajuški net*}
       lit. 'To.us$_Y$ dinner$_X$ [is] not into dinner$_{L(X)}$ if the.lady.of.the.house
       is.absent'. =
       'The dinner **cannot be enjoyed** by us if the lady of the house is absent'.

The sentence in (30a) features the syntactic idiom ⌐KAK DEREVO⌐ 'as tree'. A sign of this form can be represented as follows:
[X] ⌐КАК L(X)⌐ =
⟨'quite an ordinary [X]'; $[X_{(N)NOM}]$–**attributive**→KAK–**comparative-conjunctional**→$L(X)_{(N)NOM}$;
                    Σ = syntactic idiom, adnominal adverb, signalative; **colloquial**⟩

Its signifier is a SSynt-subtree that includes the lexeme КАК 'as' (a comparative conjunction) and a bound lexemic variable $L(X)_{(N)NOM}$, which refers to $X_{(N)NOM}$, the semantic actant of the expression; it is this variable that denotes the operation of duplication. Since the meaning of the expression cannot be distributed between its components, it is an idiom; and since its signifier is not segmental (it contains the operation of duplication), it is a syntactic idiom. Here is its lexical entry:

[X] ⌐КАК L(X)⌐, syntactic idiom, adnominal adverb, signalative; **colloquial**

Definition
'[X] ⌐КАК L(X)⌐' = 'quite an ordinary X'

Government Pattern

$X \Leftrightarrow I$
1. N_{NOM}

Ja prosto čelovek$_X$ *kak čelovek*$_{L(X)}$ lit. 'I [am] simply a.human as a.human'. = 'I am simply quite an ordinary human'.

Surface-Syntactic Implementation[12]
1. Ξ−**subjectival**→[X]−**attributive**→KAK−**compar-conjunctional**→L(X) | **no**
    X−**modificative**→Ψ
    Pered nami stojal$_\Xi$ *dom*$_X$−**attr**→*kak dom*$_{L(X)}$ lit. 'Before us [there] stood house as house'. =
    'Before us there was quite an ordinary house'.
2. [X]←**subjectival**−BYT′−**copular-completive**→KAK−**compar-conjunctional**→L(X)|**no**
    X−**modificative**→Ψ
    Dom$_X$ *ètot byl*−**cop-compl**→*kak dom*$_{L(X)}$ lit. 'House this was as house'. =
    'This house was quite ordinary'.

The sign presented in (30b) also has as its signifier a SSynt-subtree that contains three genuine lexemes—BYT′**I.3** 'be', NE 'not' and V 'in'—and a bound lexemic

---

12 These SSynt-specifications express the following linguistic facts.
In the SSyntS, the subtree implementing the Russian syntactic idiom [X] ⌐КАК L(X)⌐ can function in one of two SSynt-roles:
- either as an attribute of the SSynt-subject—that is, the semantic actant X of the idiom must be implemented as the subject of the clause;
- or as the copular complement of the verb BYT′ 'be', which has X as the subject.
Moreover, X cannot have a modifier: **Èto byl kamennyj dom kak dom* 'This was quite an ordinary stone house'.

variable L(X), to be filled with a duplicate of the actant X; thus, this sign is non-segmental. It is another syntactic idiom:

[Y-*u*] [X] ⌐BYT′I.3 NE V L(X)⌐, syntactic idiom, full clause, signalative; **colloquial**

Definition
'[Y-*u*] [X] ⌐BYT′**I.3** NE V L(X)⌐' = 'X cannot be enjoyed by Y'

Government Pattern

X ⇔ **I**	Y ⇔ **II**
1. N$_{NOM}$	1. N$_{DAT}$

Sejčas detjam$_Y$ kanikuly$_X$ ne v kanikuly
lit. 'Now to.kids$_Y$ holidays$_X$ are not into holidays'. = 'Now holidays cannot be enjoyed by kids'.

Surface-Syntactic Implementation

```
     ┌───indir-objectival───┐┌───copul-completive───┐
     ▼                      ▼▼                       ▼
[Y] [X]←subjectival−BYT′I.3 NE←restrictive−V−prepositional→L(X)
```

The ground is ready for the formal definition of syntactic idiom.

3.3 Syntactic idioms

3.3.1 The definition of syntactic idiom

Since idioms are a type of phraseme, in order to define syntactic idioms, syntactic phrasemes must first be defined.

The syntactic phrasemes, just like lexemic and morphemic phrasemes, are a major subclass of phrasemes. However, unlike the latter, syntactic phrasemes are *non-segmental signs*. In prose, this means that they do not consist exclusively of lexemes or morphemes: the signifier of a syntactic phraseme includes prosody (intonations, pauses, accents) or an operation (for instance, the duplication of an actant).

Definition 5: syntactic phraseme (see Mel'čuk 1987)

Let there be a complex sign **s** = ⟨'σ' ; /s/ ; Σ$_s$⟩ involving at last two minimal syntactic subtrees (of the form L$_1$−**synt**→L$_2$) which is constrained: that is, **s** is such that none of its components can be selected freely by the Speaker to express the meaning 'σ'; cf. Definition 1.

> A constrained complex sign **s** is a *syntactic phraseme* if and only if its signifier /s/ is non-segmental, that is, /s/ contains prosody or a bound lexemic variable, e.g., L(X), symbolizing the operation of duplication of the phraseme's actant X.

> **NB** The genuine lexemes that appear as components of a syntactic phraseme belong mostly to closed lexical-syntactic classes: they are prepositions, conjunctions or particles.

Remember, the signifier of a lexemic phraseme is segmental: such a phraseme contains only genuine lexemes of the language, but no special prosody and no bound lexemic variables. The difference between a lexemic phraseme and a syntactic phraseme, whose signifier is non-segmental—it contains special prosody or a bound lexemic variable—is obvious.

For the time being, only one class of syntactic phrasemes is known: *syntactic idioms*.

Definition 6: syntactic idiom

> A syntactic phraseme is a *syntactic idiom* if and only if it is non-compositional.

It is not known whether compositional syntactic phrasemes—that is, syntactic collocations and syntactic clichés—are possible. Therefore, in what follows I will speak exclusively of syntactic idioms.

All the syntactic phrasemes cited above are syntactic idioms, since the meaning of each of them cannot be regularly constructed by uniting the meanings of its components.

In conformity with Definitions 5 and 6, syntactic idioms come in two structural types.

- **Syntactic idioms whose signifier contains prosody**
  - Idioms containing no lexemes at all, as, for instance, ⌜[*Petja*$_X$] [*u nas*$_Z$] [*budet guljat'*$_Y$ *po nočam!*]⌝ lit. 'Pete$_X$ at us$_Z$ will.go.for.a.walk$_Y$ at night!' = 'If Pete goes for a walk at night, he will be severely punished by us'. All lexemes of the sentence *Petja u nas budet guljat' po nočam!* are part of the actants of the idiom (that is, the actants of the predicate 'punish'). The meaning of the idiom (a threat of punishment: 'If X does Y, X will be punished by Z for Y(X)') is expressed, as explained above, by the prosody imposed on the string of its actants.
  - Idioms containing just one lexeme (plus prosody), as, for instance, čтобы 'that' in ⌜Čtoby [*Petja guljal*$_X$ *po nočam*]?⌝ lit. 'That Pete go.for.a.walk$_X$ at night?' The clause *Petja guljaet*$_X$ *po nočam* is the actant of the idiom ('It is quite impossible [that X takes place]'), while the conjunction čтобы 'that' is a component of it, even if it contributes nothing to its meaning,

a sarcastic negation: 'quite impossible for Pete to go for a walk at night'. This meaning is, again, expressed by prosody.

- **Syntactic idioms whose signifier contains a bound lexemic variable** L, which can be one of the two types:
    - Either L is filled with any lexical expression that, however, is semantically bound to actant X: it must denote an intensifier of X (see (28a), p. 50);
    - or L is filled with a duplicate of the idiom's actant: [*Mal'čiki$_X$*] ⌐*est' mal'čiki$_{L(X)}$*⌐ 'Boys are boys' (see all other cases above). Syntactic idioms of the last type are the most numerous among the syntactic idioms known to me (in any case, in Russian).

3.3.2 Distinguish and avoid confounding!

Since the notion of syntactic idiom is relatively new and not stable enough, it is worth insisting on the differences between syntactic idioms and the following four other types of phrase that look similar enough to, and maybe confounded with, the former.

1) A phrase describable by a semantically loaded (= meaningful) SSynt-relation. Thus, the well-known Russian approximate-quantitative construction of the type *knig pjatnadcat'* lit. 'books fifteen' = 'maybe fifteen books' manifests the **approximate-quantitative** SSynt-relation:

PJATNADCAT'←**approx-quantitative**−KNIGA$_{PL}$ ⇔ *knig pjatnadcat'* 'maybe fifteen books'

The ordinary quantitative construction is implemented by a phrase with the inverse word order and a different SSynt-relation:

PJATNADCAT'←**quantitative**−KNIGA$_{PL}$ ⇔ *pjatnadcat' knig* 'fifteen books'

The **approximate-quantitative** SSyntRel is meaningful (contrary to the **quantitative** SSyntRel, which does not carry any meaning): it contributes to the meaning of the phrase the semantic component 'the Speaker is not sure of the number he indicates'. In our description, this component can be encoded in the deep-syntactic structure by means of the fictitious lexeme «PRIMERNO» 'maybe'.[13] It is

---

13 A fictitious lexeme can formally coincide with a real lexeme of the language. Thus, along with the fictitious «PRIMERNO», Russian has the "genuine" lexeme PRIMERNO 'approximately', which has a meaning different from that of «PRIMERNO». The genuine PRIMERNO can combine with the approximate-quantitative construction: *knig primerno pjatnadcat'* 'maybe approximately fifteen books'.

«PRIMERNO» that appears in the deep-syntactic structure to represent the approximate-quantitative phrase; the DSynt-structure of the phrase *knig pjatnadcat'* is as follows:

$$\text{«PRIMERNO»} \leftarrow \textbf{ATTR} - \text{PJATNADCAT'} \leftarrow \textbf{ATTR} - \text{KNIGA}_{PL}$$

As another example, consider phrases of the type {*Ivan –*} *durak durakom* lit. {Ivan [is]} fool by.fool' = '{Ivan is} merely a fool' = 'Ivan is a fool times two', {*Èto*} *voda vodoj* lit. '{This [is]} water by.water' = '{This [is]} merely water', etc. They are described, on the DSynt-level, by the fictitious lexeme «PROSTO» = «MERELY» and at the SSynt-level, by means of the **"merely"-reduplicative** SSyntRel:

$$\text{«MERELY»} \leftarrow \textbf{ATTR} - \text{DURAK}_{SG} \Leftrightarrow \text{DURAK}_{SG} - \textbf{"merely"-reduplicative} \rightarrow \text{DURAK}_{SG}$$
(for more on this construction, see Janda *et al.* 2020).[14]

2) A phrase containing an ordinary lexical unit L (a lexemic idiom or a single lexeme) and L's actants. Thus, the sentence *U Ivana ruki češutsja vzjat'sja za kist'* lit. 'At Ivan hands itch to take up a.paint.brush' = 'Ivan has a strong wish to start painting' contains quite an ordinary lexemic idiom [*u* X-*a*] ⌈RUKI ČEŠUTSJA⌉ [Y-*it'*]; X and Y are the idiom's actants, and by no means "its variable lexical parts," as one sees them called from time to time. In an analogous way, in the expression ⌈K ČËRTU⌉ [Y]! lit. 'To devil [with.Y]!' Y is the actant of this lexemic idiom rather than its mystical "variable lexical part." In the sentence *Doloj ètix islamofašistov!* 'Down with these Islamofascists!' the phrase *ètix islamofašistov* expresses the actant of the lexeme DOLOJ [Y]! 'Down [with.Y]!': no idiom at all here. Similarly, the phrase *Xren tebe!* lit. 'Horseradish to.you!' = 'You won't get anything!' contains the single lexeme XREN [Y-*u*]!) and its actant *tebe* 'to.you'.

---

14 Russian has a whole family of different ...**-reduplicative** SSyntRels, necessary, in particular, for the description of such phrases as the following ones:

krasnyj-krasnyj lit. 'red-red' = 'very red'

krasnee krasnogo lit. 'more.red than.red' = 'very-very red'

krepko-prekrepko lit. 'strongly-overstrongly' = 'very-very strongly'

Kto, kto? 'Who, who?'

kto-kto[Y, a...] lit. 'who-who does Y, but ...' = 'maybe somebody does Y, but ...'

durak durakom lit. 'fool by.fool' = 'merely a fool'

{*Ivan*} *el, el,* {*i ne mog ostanovit'sja*} lit. '{Ivan} was.eating, was.eating {and was unable to stop}'.
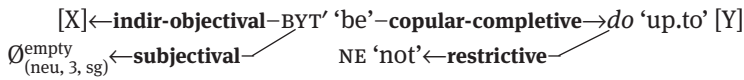
{*Sup*} *kipel-kipel,* {*da i vykipel*} lit. '{Soup} was.boiling-was.boiling, {and then boiled off}'.

Vstrečat'sja {*s nim ja*} *vstrečalsja* lit. 'To.meet {him I} met'. = 'As for meeting [him], I met him'.

3) An ordinary lexemic idiom that consists of "semi"-grammatical, or "light-weight," lexemes. Such is, for instance, the Russian lexemic idiom (31), whose lexemic components are BYT' 'be' and NE 'not':

(31) [X-*u*] ⌐BYT' NE⌐ [*do* Y-*a*] lit. '[To.X] [it] is not [up.to Y]' =
     'X cannot think of Y because X is preoccupied with something else'.
     Maše$_X$ *bylo ne do knig*$_Y$ lit. 'To.Masha [it] was not up.to books'.=
     'Masha could not think of books because she was preoccupied with something else'.

Its SSynt-structure is as follows:

$$[X]\leftarrow\textbf{indir-objectival}-\text{BYT' 'be'}-\textbf{copular-completive}\rightarrow do\text{ 'up.to' }[Y]$$
$$\emptyset^{\text{empty}}_{(\text{neu, 3, sg})}\leftarrow\textbf{subjectival}-\qquad\text{NE 'not'}\leftarrow\textbf{restrictive}-$$

4) An ordinary lexemic idiom that has some syntactic peculiarities fairly often is also called "syntactic phraseme/idiom." As an example, consider Russian lexemic idioms ⌐VSË RAVNO⌐ ≈ 'just the same' (described in Iomdin 2010: 156–162):

⌐VSË RAVNO⌐**1:** ≈ 'independently from anything' ≈ 'anyway'
Ja vsë ravno sižu doma 'I stay at home anyway'.

⌐VSË RAVNO⌐**2:** ≈ 'be.indifferent to Y'
Mne bylo vsë ravno, kuda idti 'It was the same to me where to go'.

⌐VSË RAVNO⌐**3:** ≈ 'the same as...'
Dejstvovat' tak – èto vsë ravno, čto priznat' poraženie
'To act in this way is the same as to accept defeat'.

⌐VSË RAVNO⌐**4:** ≈ 'no matter...'
Vsë ravno gde, vsë ravno kem, no èto otkrytie budet sdelano
'No matter where, no matter by whom, but this discovery will be done'.

The four idioms have the same SSynt-structure: ⌐VSË←**restrictive**–RAVNO⌐. Their syntactic peculiarities:

⌐VSË RAVNO⌐**1**   is a sentential adverb and cannot have syntactic dependents.

⌐VSË RAVNO⌐**2**   appears as a copular complement and governs—via the copula verb BYT' 'be'—a subject subordinate clause (with the complementizer čTO 'that' or with a relative-interrogative pronoun) and an indirect object in the dative.

⌐VSË RAVNO⌐**3**   is also used as the complement of the copula BYT' 'be' and governs an oblique object (of the form čTO/KAK ≈ 'as' + N/V$_{\text{INF}}$) or

an object subordinate clause introduced by the complex conjunction ⌐ᴋᴀᴋ ᴇsʟɪ ʙʏ¬ 'as if'.

⌐ᴠsË ʀᴀᴠɴᴏ¬**4** is a component of a dozen indefinite pronouns, such as ⌐ᴠsË ʀᴀᴠɴᴏ¬ ᴋᴛᴏ 'no matter who', ⌐ᴠsË ʀᴀᴠɴᴏ¬ ᴏᴛᴋᴜᴅᴀ 'no matter where from' and ⌐ᴠsË ʀᴀᴠɴᴏ¬ ᴘᴏČᴇᴍᴜ 'no matter why' (similar to ᴋᴏᴇ-ᴋᴛᴏ 'somebody', ᴋᴛᴏ ᴜɢᴏᴅɴᴏ 'anybody', ᴍᴀʟᴏ ᴋᴛᴏ 'few people', etc.).

The four ⌐ᴠsË ʀᴀᴠɴᴏ¬ expressions are ordinary lexemic idioms featuring some syntactic particularities.

3.3.3 Fictitious lexemes

Strictly speaking, the topic of fictitious lexemes lies outside of the general discussion of phraseology. However, since fictitious lexemes are necessary to represent –both in the lexicon and in deep-syntactic structures–syntactic idioms whose signifier is only prosody, a few words have to be said about them.

Definition 7: fictitious lexeme (Mel'čuk 2013: 37–41, 2018)

> A *fictitious lexeme* of language **L** is a linguistic sign of **L** whose signified is similar to lexical signifieds of **L**, but whose signifier is non-segmental, that is, it is not a string of phonemes: it can be word order, prosody, or operations (presented by lexemic variables).

A fictitious lexeme is given a conventional name shown by « » quotes; under this name it is stored in the lexicon and appears in DSynt-structures.

> **NB** Strictly speaking, a fictitious lexeme is a type of **deep lexeme**; it is as different from a surface lexeme as a fictitious person is different from a real person.

Fictitious lexemes are used to represent linguistic phenomena of the three following types:

- Those semantic differences between phrases consisting of the same lexemes that cannot be expressed in the DSynt-structure only by means of DSynt-relations. Such is, for instance, the case of phrases with the **subject-copredicative** and **object-copredicative** SSynt-relations. These SSyntRels both correspond to the DSynt-relation **ATTR**, but carry different meanings, since their dependents semantically bear on different clause elements:

(32)   ┌──**subj-copred**──┐
                      ▼
*John knew Mary an old **man*** [John was an old man].
    vs.   ┌──**obj-copred**──┐
                             ▼
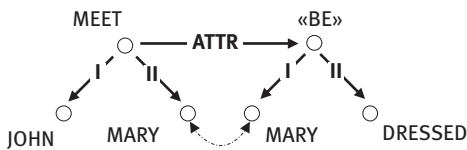*John knew Mary an old **woman*** [Mary was an old woman].

To express this semantic difference on the DSynt-level, the fictitious lexeme «BE» is introduced into the DSyntS as an **ATTR**-dependent of the Main Verb:

(33) a. [*John*] *met*–[*Mary*]–**subj-copredicative**→*dressed* [John was dressed]. ⇔

MEET — **ATTR** → «BE»
I II   I II
JOHN   MARY   JOHN   DRESSED
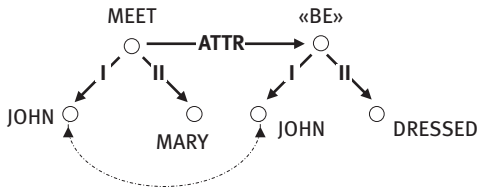
vs.

b. [*John*] *met*–[*Mary*]–**obj-copredicative**→*dressed* [Mary was dressed] ⇔

MEET — **ATTR** → «BE»
I II   I II
JOHN   MARY   MARY   DRESSED

☞ A dashed double-headed arrow indicates coreference.

- Semantically loaded (= meaningful) surface-syntactic relations, which are illustrated in 3.3.2, Item 1), with the **approximate-quantitative** SSyntRel. Meaningful SSyntRels are not widespread in English, but there are some; here is an example:

(34) *Politics, schmolitics! | Theory, schmeory. | Books, schmooks. | Baby, schmaby…*

The so-called *schm*-reduplication of a noun $L_{(N)}$ expresses the Speaker's derision and skepticism about $L_{(N)}$'s referent: '*SCHM*-$L_{(N)}$' ≈ 'I signal that I dismiss $L_{(N)}$ as being ludicrous and worthless'. In the DSyntS of a sentence with the lexeme *schm*-$L_{(N)}$, this meaning can be represented by the fictitious lexeme «DERISION», and English has the following deep-syntactic rule:

(35) $L_{(N)}$–**ATTR**→«DERISION»   ⇔ $L_{(N)}$–**schm-reduplicative**→$L_{(N)}$

Note that the rule (35) actually is a part of the lexical entry for the fictitious lexeme «DERISION»: it tells us how to implement this DSynt-node in the surface-syntactic structure. On the next step, the **schm-reduplicative** SSyntRel is expressed as *SCHM*-$L_{(N)}$, where *SCHM*- is the name of the corresponding derivational means (adding to $L_{(N)}$ the prefix /šm/- and deleting the initial prevocalic cluster in $L_{(N)}$, if any).

Russian is rich in semantically-loaded SSyntRels; here are two more examples of such SSyntRels and of their encoding in the DSynt-structure.

(36) «IF.ONLY» for the **irreal-subjectival** SSyntRel:

      DSyntS: «IF.ONLY»−**II**→ZAMETIT' 'notice'−**I**→IVAN ⇔

      SSyntS: ZAMETIT'−**irreal-subjectival**→IVAN

              Zamet'$_{\text{IMPER, 2, SG}}$ *Ivan ètu jamu, ...* 'If only Ivan had noticed this hole, ...'

(37) «AS.FOR» for the **focalizing-reduplicative** SSyntRel (Russian phrases of the corresponding type are described in detail in Iomdin 2013):[15]

            ┌──────────────**ATTR**──────────────┐

      DSyntS: «AS.FOR»−**II**→PROČEST' NE←**ATTR**−PROČEST'−**II**→ROMAN$_{\text{SG}}$ ⇔

           ┌──────────**focalizing-reduplicative**──────────┐

      SSyntS: PROČEST'−**dir-obj**→ROMAN$_{\text{SG}}$ NE←**restr**−PROČEST'

           Pročest'$_{\text{INF}}$ *roman ja ne pročla, a tol'ko perelistala*

           'As for RÉADING the novel, I did not read it, but only leafed through it'.

- Syntactic idioms whose signifier is exclusively prosody (it is because of these idioms that the notion of fictitious lexeme had to be introduced in this paper).

One such idiom—«WILL.BE.PUNISHED»—was described in Subsection 2.2, here is another one (the conversive of the former)—«WILL.PUNISH»—in sentence (38):

(38) *Ivan tebe poguljaet po nočam!* lit. 'Ivan to.you will.go.for.a.walk at night!' = 'If you go for a walk at night, Ivan will severely punish you'.

The DSynt- and SSynt-structures of this idiom taken together with its actants are as follows:

                ┌─**III**──────────┐

DSyntS: [IVAN]←**I**−«WILL.PUNISH»−**II**→[TY] [POGULJAT'] PO NOČAM   ⇔

SSyntS: [IVAN]←**subjectival**−[POGULJAT']−**indir-objectival**→[TY] PO NOČAM

It is on this structure that the idiom's prosody is imposed.

    As already mentioned, each syntactic idiom has its own lexical entry, just like lexemic idioms do; this entry gives, among other things, the information necessary to expand the DSynt-node of the idiom into its SSynt-subtree (and then to properly implement this subtree in the morphological string).

    Here is a list of fictitious lexemes gleaned from various languages (see also Mel'čuk 2013: 37–42); to facilitate the reader's task, they are represented here by their English equivalents. The fictitious lexemes feature rather abstract meanings that one could loosely qualify as "grammatical"; that is why Žolkovskij (1971: 10) spoke of them as "belonging to a border zone between the lexicon and the grammar."

---

15 The "additional" information carried by the **focalizing-reduplicative** SSyntRel is the indication that its dependent element implements the emphatic theme of the clause.

NB In the strict sense of the term, a fictitious lexeme is not *a grammatical lexeme*: it carries a specific meaning, it is not introduced into the SSyntS by syntactic rules, it does not express an inflectional value, and it is not a pronoun. But it is similar to a grammatical lexeme by its broad and vague enough meaning.

«AFFECT»	«CAUSE$_{(N)}$»	«IF»	«MOVE$_{DIR}$»
«AFTER»	«CONDITION»	«IF.ONLY»	«NAME»
«AS.FOR»	«DERISION»	«INCLUDE»	«NUMBER [*of*]»
«BE»	«FOR» (*buy **her** a dress*)	«INSTRUMENT»	«SAY»
«BE.ABLE»	«FROM» (*one **of** these*)	«MATERIAL»	«SHOULD»
«BE.FROM»	«GOAL»	«MAYBE»	«TITLE» (***Professor** Drouin*)
«BECOME»	«HAVE»	«MERELY»	«WHILE»
«BELONG»	«HAVE.TO»	«MORE»	«WITH»

NB This list does not include the fictitious lexemes postulated for the Russian purely pro-sodic syntactic idioms in this paper.

3.3.4 An illustrative list of Russian syntactic idioms

Syntactic idioms include lexemes mainly from closed lexical classes or are implemented only by prosody; therefore, they cannot be very numerous. However, they are of special interest from a theoretical angle, so that it is worth presenting here those Russian idioms that are known to me at present.

NB 1. This list is, of course, far from complete; for a more detailed enumeration of hypothetical Russian syntactic idioms, see Kopotev (2008: 124–126).
    2. The syntactic idioms in the list below are not described lexicographically, and their meanings are formulated approximately; neither their SSynt-structures, nor their prosodies are shown (while, for instance, the idioms Nos. 11, 12 and 13 are distinguished only by prosody—example (29), p. 52).

1. ⌜BYT′**I.3** VSEM L(X)-*am*⌝ [X] 'X is the most outstanding of Xs'
   {*Èto*} *vsem borščam*$_{L(x)}$ *boršč*$_x$*!* lit. '{This} [is] to.all borschts$_{L(x)}$ borscht$_x$!' = 'This is the most outstanding of all possible borschts!'

2. ⌜ČTO [X], TO L(X)⌝ 'I admit that this is really an X'
   Čto Ivan durak$_x$*, to durak*$_{L(x)}$ lit. 'What Ivan [is] fool$_x$, so fool$_{L(x)}$'. = 'I admit that Ivan is really a fool'.
   Čto xolodno$_x$*, to xolodno*$_{L(x)}$ lit. 'What [is] cold$_x$, so cold$_{L(x)}$'. = 'I admit that it is really cold'.

   NB Detailed descriptions of the syntactic idioms Nos. 2, 23 and 28 are found in Kopotev 2005 and Kopotev and Fajnvejc 2007.

11. «WILL.BE.COERCED» = $\ulcorner$[X$_{(N)NOM}$] [u Z$_{(N)GEN}$] [Y$_{(V)FUT}$]!$\urcorner$ [**assurance**]
      'X will be coerced to do Y by Z'
    Ivan$_X$ u nas$_Z$ vyučit$_Y$ francuzskij! lit. 'Ivan$_X$ at us$_Z$ will learn$_Y$ French!' =
    'With us Ivan will learn French all right!'

12. «WILL.NOT.BE.ALLOWED» = $\ulcorner$[X$_{(N)NOM}$] [u Z$_{(N)GEN}$] [Y$_{(V)FUT}$]$\urcorner$ [**sarcastic negation**]
      'X will not be allowed to do Y by Z'
    Ivan$_X$ u nas$_Z$ poguljaet$_Y$ po nočam lit. 'Ivan$_X$ at us$_Z$ will.go.out.for.a.walk$_Y$ at night'. =
    'We won't let Ivan go out for a walk at night'.

13. «WILL.BE.PUNISHED» = $\ulcorner$[X$_{(N)NOM}$] [u Z$_{(N)GEN}$] [Y$_{(V)FUT}$]!$\urcorner$ [**threat**] 'X will be
    punished for Y by Z'
    Ivan$_X$ u nas$_Z$ poguljaet$_Y$ po nočam! lit. 'Ivan$_X$ at us$_Z$ will.go.for.a.walk$_Y$ at night!' =
    'If Ivan goes for a walk at night, he'll be severely punished by us!'

14. «WILL.PUNISH» = $\ulcorner$[X$_{(N)NOM}$] [Z$_{(N)DAT}$] [Y$_{(V)FUT}$]!$\urcorner$ [**threat**] 'X will punish Z for Y'
    Ivan$_X$ tebe$_Z$ poguljaet$_Y$ po nočam! lit. 'Ivan$_X$ to you$_Z$ will.go.for.a.walk$_Y$ at night!' =
    'If you go for a walk at night, Ivan will severely punish you!'

> **NB** The syntactic idioms «WILL.BE.PUNISHED» and «WILL.PUNISH» are conversives:
>   *On u nas poguljaet po nočam!* ≡ *My emu poguljaem po nočam!*

15. [X] $\ulcorner$(BYT′**I.2**) KAK L(X)$\urcorner$ '(be) quite an ordinary X'
    Èto byla komnata$_X$ kak komnata$_{L(x)}$ lit. 'This was room$_X$ as room$_{L(x)}$'. =
    'This was quite an ordinary room'.
    Komnata$_X$ byla kak komnata$_{L(x)}$ lit. 'Room$_X$ was as room$_{L(x)}$'. = 'This room was
    quite ordinary'.

16. [X] $\ulcorner$BYT′**I.2** L(X)$\urcorner$ 'X has well-known properties'[17] (Vilinbaxova and Kopotev 2017.)
    {*Nu,*} *mužčiny$_X$ est′ mužčiny$_{L(x)}$* '{Well,} men$_X$ are [lit. 'is'] men$_{L(x)}$'. =
    'Everybody knows what can be expected of men'.

---

17 Idioms 16 and 17 are tautology-like; and as such, they are a hot topic. Their meanings are quite vague, so that their interpretation by the Addressee strongly depends on the context. As a result, linguists did not reach a consensus about their lexicographic definitions (see Vilinbaxova and Kopotev 2017 for a discussion and relevant references). The main differences between idioms 16 and 17 are, in my opinion, as follows.

 – Semantically, [X] $\ulcorner$BYT′**I.2** L(X)$\urcorner$ (No. 16: *Vojna est′ vojna* lit. 'War is war') means that Xs have well-known properties of which the Addressee must be aware. (These "well-known properties" are, in most cases, the *lexicographic connotations* of the noun X: see Apresjan 1995: 166–167.) But [X] – $\ulcorner$ÈTO BYT′**I.3** L(X)$\urcorner$ (No. 17: *Vojna – èto vojna* lit. 'War—it [is] war') means that the word X is used in its exact sense, so that it means precisely what it says, and the Addressee is supposed to know this.

 – Syntactically, in [X] $\ulcorner$BYT′**I.2** L(X)$\urcorner$ (No. 16), X can be only a noun; but [X] – $\ulcorner$ÈTO BYT′**I.3** L(X)$\urcorner$ (No. 17) accepts as its actant X a lexical expression of any type (a noun, an adverb, a prepositional phrase, a verb, etc.).

17. [X] ⌜– ÈTO BYT′**I.3** L(X)⌝ 'X is exactly an X' (Vilinbaxova and Kopotev 2017.)
    {*Nu,*} *mužčina*$_x$ *– èto mužčina*$_{L(x)}$ lit. '{Well,} man$_x$—this [is] man$_{L(x)}$'. =
    'A man is exactly a man'.
    Blizko$_x$ *– èto blizko*$_{L(x)}$ 'Close $_x$—this [is] close$_{L(x)}$'. = 'Close is exactly close'.
    Uexala$_x$ *– èto uexala*$_{(x)}$ '[She].left $_x$—this [is] [she].left$_{L(x)}$'. = 'She left is exactly
    she left'.

18. [X] ⌜I L(X)⌝ 'X is nothing special'
    {*A čto takogo?*} *Vyšel*$_x$ *i vyšel*$_{L(x)}$ lit. '{And what is there?} [He] went.out$_x$ and
    went.out$_{L(x)}$'. =
    'He went out, and this is nothing special'.
    {*Nu,*} *gost′*$_x$ *i gost′*$_{L(x)}$ lit. '{Well,} guest$_x$ and guest$_{L(x)}$'. =
    'Well, there is a guest, and this is nothing special'.

19. [X] ⌜L(X)-*om,* A⌝ [Y] 'Let's leave X out of discussion because Y'
    Vodka$_x$ *vodkoj*$_{L(x)}$, *a rabota ne ždët*$_Y$*!* lit. 'Vodka$_x$ by.vodka$_{L(x)}$, but work does
    not wait$_Y$!' =
    'Let's leave vodka out, since we have to work'.

20. [X] ⌜L(X)-*om,* A [Y] L(Y)-*om*⌝
    'Both X and Y are important and should not interfere with each other'.
    Rabota$_x$ *rabotoj*$_{L(x)}$, *a obed*$_Y$ *obedom*$_{L(Y)}$ lit. 'Work$_x$ by.work$_{L(x)}$, but lunch$_Y$
    by.lunch$_{L(Y)}$'. =
    'Both work and lunch are important and should not interfere with each other'.

21. [X] ⌜*na* L(X)-*e* SIDIT I L(X)-*om* POGONJAET⌝ 'Xs are everywhere'
    Zdes′ žulik$_x$ *na žulike*$_{L(x)}$ *sidit i žulikom*$_{L(x)}$ *pogonjaet*
    lit. 'Here crook$_x$ is.sitting on crook$_{L(x)}$ and is.driving with.crook$_{L(x)}$'. =
    'Here there are crooks everywhere'.

22. [X] ⌜NE L(X)⌝ 'This is not quite an X'
    Mordoboj$_x$ *ne mordoboj*$_{L(x)}$ {, *prosto spor*} lit. 'Tussle$_x$ not tussle$_{L(x)}$ {, simply
    discussion}'. =
    'This is not quite a fight, this is simply a discussion'.

23. [X] ⌜TAK L(X)⌝ 'I accept X'[18]
    Piva$_x$ *tak piva*$_{L(x)}$*!* lit. 'Of.beer$_x$ so of.beer$_{L(x)}$!' = 'OK, I'll have the beer'.
    Ždat′$_x$ *tak ždat′* lit. 'To.wait so to.wait'. = 'OK, I'll wait'.
    Po-xorošemu$_x$ *tak po-xorošemu* lit. 'Nicely so nicely'. = 'OK, nicely then'.

---

18 This idiom is used only in a dialog—as a reaction to the previous utterance of the Addressee.
The actant X must coincide formally—that is, phoneme by phoneme—with the Addressee's word-
form with whose referent the Speaker agrees. The same is true of the syntactic idiom No. 24.

24. $[X]^{\lceil}$-TO $L(X)^{\rceil}$ 'As far as X is concerned, it is X'
Po forme$_X$-*to* {, *konečno,*} *po forme*$_{L(x)}$. {*Odnako delo ne v ètom.*}
lit. 'According.to form$_X$, well, of.course, according.to form$_{L(x)}$. {The problem, however, is somewhere else'.} =
'As for being according to form, it is according to form. The problem, ...'.
Spal$_X$⟨*Spat'*$_X$⟩-*to on spal*$_{L(x)}$, *no vsë videl* lit. 'Slept$_X$⟨*To.sleep*$_X$⟩, well, he slept$_{L(x)}$, but [he] saw everything'. = 'As for sleeping, he slept, but [he] saw everything'.

25. $[X]^{\lceil}$-TO ON BYT'$_{PRES}$ $L(X)^{\rceil}$ 'As for being X, he is X'
Karlik$_X$-*to on karlik*$_{L(x)}$; *no nos u nego ogromnyj*
lit. 'Dwarf, well, he is dwarf, but his nose is enormous'. =
'As for being a dwarf, he is a dwarf; but he has an enormous nose'.

26. $[X,]^{\lceil}$XOT' $L(\text{INTENS}(X))_{(v)\text{IMPERF, IMPER, 2, SG}}^{\rceil}$ 'extremely X'[19]
{*Ivan takoj*} *xudoj*$_X$, *xot' stroenie skeleta na nëm izučaj*$_{L(\text{INTENS}(X))(v)\text{IMPERF, IMPER, 2, SG}}$
'{Ivan is} so skinny$_X$, you might even study$_{L(\text{INTENS}(X))}$ the structure of the human skeleton on him'.

27. $[X\text{-}it]^{\lceil}$ I $L(X)$-*it*$^{\rceil}$ 'continues to do X'
'*Ivan pel*$_X$ *i pel*$_{L(x)}$ 'Ivan sang$_X$ and sang$_{L(x)}$'. = 'Ivan kept singing'.
Ivanu xotelos' pet'$_X$ *i pet'*$_{L(x)}$ 'Ivan felt like singing$_X$ and singing$_{L(x)}$'. =
'Ivan felt like keeping on singing'.

28. $[X\text{-}it']^{\lceil}$ TAK $L(X)$-*it'*!$^{\rceil}$ 'Let's really do X!'
Veselit'sja$_X$ *tak veselit'sja*$_{L(x)}$! lit. 'To.have.fun$_X$ so to.have.fun$_{L(x)}$!' = 'Let's really have fun!'

29. $[Y\text{-}u]^{\lceil}$ (I) $[X]$ BYT'**I.3** NE V $L(X)^{\rceil}$ 'X cannot be enjoyed by Y'
Nam$_Y$ (*i*) *otdyx*$_X$ *ne v otdyx*$_{L(x)}$ lit. 'To.us$_Y$ even vacations$_X$ are not into vacations$_{L(x)}$'. =
'We can't even enjoy vacations'.

---

19 This Russian syntactic idiom is quite exotic in that its lexical filling $L(\text{INTENS}(X))$ is relatively free, but it must express an intensification of X. Here is another example of its use:

(i) *My byli užasno golodny*$_X$, $^{\lceil}$*xot' podošvy žuj*$^{\rceil}_{L(\text{INTENS}(X))}$!
lit. 'We were horribly hungry$_X$, even chew$_{\text{IMPERF, IMPER, 2, SG}}$ shoe.soles!' =
'..., we might even be ready to chew shoe soles!'
DSyntS: GOLODNYJ−**ATTR**→$^{\lceil}$XOT' $L(\text{INTENS}(X))_{\text{IMPERF, IMPER, 2, SG}}^{\rceil}$ ⇔
SSyntS: GOLODNYJ−**appositive**→ŽEVAT'$_{\text{IMPERF, IMPER, 2, SG}}$−**restrictive**→XOT'
             └−**direct-objectival**→PODOŠVA$_{PL}$

Thus, in this case one can say *xot' zemlju glotaj* 'you might even be ready to swallow soil' or *xot' na proxožix brosajsja* 'you might even be ready to attack passers-by', etc. To put it differently, the Speaker has the liberty to produce his own *hapax legomenon* idiom.

This list presents both structural types of syntactic idiom:
– A syntactic idiom whose signifier contains prosody, that is, which has less than two lexemes: either one or none. The idioms of this type are Nos. 3, 5, and 11–14.
– A syntactic idiom whose signifier contains a lexemic variable that is filled with the duplicate of one of the idiom's actants. The idioms of this type are all the others: Nos. 1, 2, 4, 6–10 and 15–29.

4 Conclusion

Introducing the notions of morphemic and syntactic phrasemes allows for the elaboration of a universal typology of phrasemes, which will be quite formal and strictly deductive. On the one hand, this contributes to a better description of phrasemes in the lexicon and in the grammar; on the other hand, this constitutes a step toward the construction of the coherent notional system and a formalized metalanguage for linguistics.

Acknowledgments

References

Apresjan, Valentina. 2014. Syntactic idioms across languages: Corpus evidence from Russian and English. *Russian Linguistics* 38(1). 187–203.
Aronoff, Mark & Shikaripur N. Sridhar. 1984. Agglutination and composition in Kannada verb morphology. In David Testen, Veena Mishra & Joseph Drogo (eds.), *Lexical Semantics,* 3–20. Chicago: Chicago Linguistic Society.
Avgustinova, Tania & Leonid Iomdin. 2019. Towards a typology of microsyntactic constructions. In Gloria Corpas-Pastor & Ruslan Mitkov (eds.), *Computational and Corpus-Based Phraseology*, 15–30. Cham: Springer.
Beck, David & Igor Mel'čuk. 2011. Morphological phrasemes and Totonacan verbal morphology. *Linguistics* 49(1). 175–228.

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

Čermák, František. 2007. Idioms and morphology. In Harald Burger, Dmitrij Dobrovol'skij, Peter Kühn & Neal Norrick (eds.), *Phraseology. An International Handbook of Contemporary Research*, 20–26. Berlin & New York: W. de Gruyter.

Dillon, George. 1977. *Introduction to contemporary linguistic semantics*. Englewood Cliffs, N.J.: Prentice-Hall.

Dobrovol'skij, Dmitrij, Mixail Kopotev, & Ljudmila Poppel. 2019. Gruppa konstrukcij Nu i X: semantika, pragmatika, sočetaemost' [The group of constructions of Nu ı X type: semantics. pragmatics, combinatorics]. *Scando-Slavica* 65(1). 5–25.

Goldberg, Adele. 1995. *A Construction Grammar approach to argument structure*. Chicago & London: The Chicago University Press.

Goldberg, Adele & Ray Jackendoff. 2004. English resultative as a family of constructions. *Language* 80(3). 532–568.

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

Iomdin, Leonid. 2006a. Mnogoznačnye sintaksičeskie frazemy: meždu leksikoj i sintaksisom [Polysemous syntactic phrasemes: between lexicon and syntax]. In Natal'ja Laufer, Aleksandr Narin'jani & Vladimir Selegej (eds.), *Komp'juternaja lingvistika i intellektual'nye texnologii*, 202–206. Moskva: RGGU Publisher.

Iomdin, Leonid. 2006b. Novye nabljudenija nad sintaksisom russkix frazem [New observations on the syntax of Russian phrasemes]. In Bożena Chodźko, Elżbieta Feliksiak, Marek Olesiewicz (eds.), *Obecność*, 247–281. Białystok: Uniwersytet w Białymstoku.

Iomdin, Leonid. 2008. V glubinax mikrosintaksisa: odin leksičeskij klass sintaksičeskix frazem [In the depths of microsyntax: a lexical class of syntactic phrasemes]. *Komp'juternaja lingvistika i intellektual'nye texnologii* 7(14).178–184.

Iomdin, Leonid. 2010. Sintaksičeskie frazemy: meždu leksikoj i sintaksisom [Syntactic phrasemes: between lexicon and syntax]. In Jurij Apresjan, Igor' Boguslavskij, Leonid Iomdin & Vladimir Sannikov, *Teoretičeskie problemy russkogo sintaksisa. Vzaimodejstvie grammatiki i slovarja*, 141–190. Moskva: Jazyki slavjanskix kul'tur.

Iomdin, Leonid. 2013. *Čitat' ne čital, no*…: ob odnoj russkoj konstrukcii s povtorjajuščimisja slovesnymi èlementami [*Čitat' ne čital, no*…: on a Russian construction with repeated lexical elements]. *Komp'juternaja lingvistika i intellektual'nye texnologii* 12(19). 297–310.

Iomdin, Leonid. 2017. Meždu sintaksičeskoj frazemoj i sintaksičeskoj konstrukciej. Netrivial'nye slučai mikrosintaksičeskoj neodnoznačnosti [Between the syntactic phraseme and the syntactic construction. Nontrivial cases of microsyntactic ambiguity]. *Slavia* 68(2–3). 230–243.

Jackendoff, Ray. 2008. Construction after construction and its theoretical challenges. *Language* 84(1). 8–28.

Janda, Laura, Mihail Kopotev & Tore Nesset. 2020. Constructions, their families and their neighborhoods: the case of *durak durakom* 'a fool times two'. *Russian Linguistics* 44 (1). 109–127.

Kopotev, Mixail. 2005. ГУЛЯТЬ ТАК ГУЛЯТЬ: Between grammar and dictionary. In Jurij Apresjan & Leonid Iomdin (eds.), *East West Encounter. Second International Conference on Meaning – Text Theory*, 225–236. Moscow: Slavic Culture Languages Publishing House.

Kopotev, Mixail. 2008. *Principy sintaksičeskoj idiomatizacii* [Principles of syntactic idiomatization]. Helsinki: Helsinki University Press.

Kopotev, Mixail & Aleksandra Fajnvejc. 2007. *Izučat' tak izučat'*: sinxronija i diaxronija [*Izučat' tak izučat'*: synchrony and diachrony]. *Naučno-texničeskaja informacija, serija 2. Informacionnye processy i sistemy* 9. 103–123.

Kopotev, Mixail & Tat'jana Steksova. 2016. *Isključenie kak pravilo: perexodnye edinicy v grammatike i slovare* [The exception as a rule: transitional units in grammar and in lexicon]. Moskva: Jazyki slavjanskix kul'tur.

Mel'čuk, Igor'. 1964. Obobščenie ponjatija frazeologizma (morfologičeskie "frazeologizmy") [Generalizing the notion of phraseologism (morphological "phraseologisms")]. In Leonid 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"], vol. I, Samarkandskij Gosudarstvennyj Universitet: Samarkand, 89–90.

Mel'čuk, Igor. 1987. Un affixe dérivationnel et un phrasème syntaxique du russe moderne. Essai de description formelle. *Revue des études slaves*, 59(3). 631–648.

Mel'čuk, Igor. 1993–2000. *Cours de morphologie générale, vols. 1–5.* Montréal, Les Presses de l'Université de Montréal & Paris: CNRS Éditions.

Mel'čuk, Igor. 1995. Phrasemes in language and phraseology in linguistics. In Martin Everaert, Erik-Jan van der Linden, André Schenk & Rob Schreuder (eds.), *Idioms. Structural and Psychological Perspectives*, 167–232. Hillsdale, N.J. & Hove, U.K.: Lawrence Erlbaum Associates.

Mel'čuk, Igor. 2001. *Communicative organization in natural language. The semantic-communicative structure of sentences*. Amsterdam & Philadelphia: John Benjamins.

Mel'čuk, Igor. 2004. La non-compositionnalité en morphologie linguistique. *Verbum* 26(4). 439–458. See also: http://olst.ling.umontreal.ca/pdf/CompositionnaliteEnMorphologie.pdf.

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

Mel'čuk, Igor. 2008. Zero affixes and nominal cases in Daghestanian languages. In A. Arxipov, L. Zaxarov, A.A. Kibrik, A.E. Kibrik, I. Kobozeva, O. Krivnova, E. Ljutikova & O. Fëdorova (eds.), *Fonetika i nefonetika. K 70-letiju Sandro V. Kodzasova*, 176–183. Moskva: Jazyki slavjanskix kul'tur.

Mel'čuk, Igor. 2012. Phraseology in the language, in the dictionary, and in the computer. *Yearbook of Phraseology* 3. 31–56.

Mel'čuk, Igor. 2013. *Semantics: From meaning to text, vol. 2.* Amsterdam & Philadelphia: John Benjamins.

Mel'čuk, Igor. 2015a. Clichés, an understudied subclass of phrasemes. *Yearbook of Phraseology* 6. 55–86.

Mel'čuk, Igor. 2015b. *Semantics: From meaning to text. Vol. 3.* Amsterdam & Philadelphia: John Benjamins.

Mel'čuk, Igor. 2018. "Wordlets": One of Zholkovsky's major contributions to the notion of deep-syntactic structure. In Dennis Ioffe, Marcus Levitt, Joe Peschio & Igor Pilshchikov (eds.), *A/Z: Essays in Honor of Alexander Zholkovsky*, 350–360. Boston: Academic Studies Press.

Mel'čuk, Igor. 2020. Clichés and pragmatemes. *Neophilologica* 32. 9–20.

Pike, Kenneth. 1961. Compound affixes in Ocaina. *Language* 37(4). 570–581.

Raxilina, Ekaterina, ed. 2010. *Grammatika konstrukcij* [Construction Grammar]. Moskva: Azbukovnik.

Rhodes, Russell. 2009. Tautological constructions in English… and beyond. https://www.semanticscholar.org/paper/Tautological-constructions-in-English…and-beyond-Rhodes/928c09fca831c1ef6ba61a5ac9f0d71c1167f99c.

Šmelëv, Dmitrij. 1960. O «svjazannyx» sintaksičeskix konstrukcijax v russkom jazyke ["Bound" syntactic constructions in Russian]. *Voprosy Jazykoznanija* 5. 47–60.

Švedova, Julija. 1960. *Očerki po sintaksisu russkoj razgovornoj reči* [Essays on the syntax of colloquial Russian]. Moskva: AN SSSR.

Švedova, Julija, ed. 1970. *Grammatika sovremennogo russkogo literaturnogo jazyka* [Grammar of modern literary Russian]. Moskva: AN SSSR.

Švedova, Julija, ed. 1980. *Russkaja grammatika. Tom II. Sintaksis* [Russian grammar. Vol. II. Syntax]. Moskva: Nauka.

Vilinbaxova, Elena & Kopotev, Mixail. 2017. «X est′ X» značit «X èto X»? Iščem otvet v sinxronii i diaxronii [«X est′ X» means «X èto X»? We are looking for an answer in synchrony and diachrony]. *Voprosy Jazykoznanija* 3. 110–124.

Weinreich, Uriel. 1969. Problems in the analysis of idioms. In Jaan Puhvel (ed.), *Substance and Structure of Language*, 23–81. Berkeley & Los Angeles: University of California Press.

Wierzbicka, Anna. 1987. *Boys will be boys* – 'Radical Semantics' vs. 'Radical Pragmatics'. *Language* 63(1). 95–114. (See also in Wierzbicka, Anna. 1991. *Cross-cultural pragmatics. The semantics of human interaction*, 391–452. Berlin & New York: Mouton de Gruyter).

Žolkovskij, Aleksandr. 1971. *Syntaksis somali* [Somali syntax]. Moskva: Nauka.