DE GRUYTER MOUTON DOI 10.1515/phras-2015-0005 — YoP 2015; 6: 55–86

Igor Mel'čuk

Clichés, an Understudied Subclass of Phrasemes

Abstract: Natural languages have three major families of phrasemes:
- Lexical phrasemes (phraseologized phrases): *kick the bucket, black box, pay a visit, if you know what I mean*.
- Morphological phrasemes (phraseologized wordforms, see [Beck & Mel'čuk 2011]): *for+get*, *light+house*, *in+dispens+able*.
- Constructional, or syntactic, phrasemes (phraseologized constructions, or phrase schemata): "$X_{(N)}$ V_{INF}?!?", as in *John be afraid?!?*

For more details on phrasemes within the Meaning-Text framework, see (Mel'čuk 2015: 336–340).

This paper proposes an overview of an important subclass of lexical phrasemes that has not been as yet paid sufficient attention: clichés.

First, a typology of lexical phrasemes is presented, with the definitions that underlie the subsequent discussion; this provides a formal framework for the description of clichés (Section 1). Second, the class of clichés is examined and a typology of clichés is proposed, based on the type of the cliché's referent (Section 2). Third, a dimension for restrictions on the use of lexemic expressions in particular situations of linguistic communication is introduced—namely, a pragmatic constraint. Applied to clichés, it defines a subclass of a subclass of clichés—pragmatemes (Section 3).

Keywords: phraseology, typology of phrasemes, idioms and collocations, clichés, pragmatemes

Igor Mel'čuk, Observatoire de linguistique Sens-Texte (OLST), Département de linguistique et de traduction, Université de Montréal, C.P. 6128, succ. Centre-ville, Montréal (Québec) H3C 3J7, Canada, E-mail: igor.melcuk@umontreal.ca

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

1 Lexical phrasemes

1.1 Introductory remarks

To examine and describe phrasemes, this paper adopts the viewpoint of linguistic synthesis: from meaning to text. In other words, within the proposed approach, the only legitimate question is "How does the Speaker produce this phraseme?", not "How does an addressee understand this phraseme?"

☛ 1. *The Speaker*—with a capital "S"—stands for 'the author of this speech act' = 'I'; *a speaker* refers to anybody speaking the language.
2. The onomasiological approach to phraseology is used in accordance with the general linguistic principle—to model natural language in the meaning-to-text direction: see (Mel'čuk 2012a, 2013, 2015).

A lexical phraseme is a non-free, or constrained, phrase, i.e., a multilexemic expression, as opposed to a free phrase. Therefore, a free phrase of language **L** has first to be defined. To do this, the linguistic synthesis needs to be briefly characterized, specifying the way the Speaker constructs any phrase. Then I will indicate the difference in the production of free *vs*. non-free phrases. This will be done using some notions and formalisms of the Meaning-Text approach; see, for instance, (Mel'čuk 2001, 2012a, 2013 and 2015). Basically, I exploit different levels of linguistic representation of utterances, more precisely the semantic structure (a semantic network) and the deep-syntactic structure (a dependency tree), as well as the conceptual representation [ConcR]: a formal description of a certain conceptualization of extralinguistic reality, geared to an eventual verbalization, but as free as possible of the specific linguistic limitations (see Mel'čuk 2001: 154–158). In this paper, the conceptual representation is invoked to describe the referents of certain phrasemes, namely clichés.

1.2 Linguistic synthesis as adopted framework

☛ *E* stands for a phrase L_1—L_2—...—L_n that consists of syntactically linked lexemes L_1, L_2, ..., L_n.

A phrase *E* is produced by the Speaker in several steps of which I will here consider the first three.

NB: No claim is made concerning the **actual** (i.e., neurological) process of speech production. Such formulations as "The Speaker constructs..." are no more than a manner of speaking: what is meant is the logical possibility of presenting things in this way. Therefore, *the Speaker* as understood here can be identified with a formal linguistic model.

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

- The Speaker constructs a conceptual representation [ConcR] of the chunk of extralinguistic reality he wants to verbalize; such a chunk may include, of course, the Speaker himself, his inner world, his intentions, etc.
- For a given ConcR, the Speaker constructs a corresponding semantic representation [SemR], which specifies the linguistic meaning of the future utterance *E*. He does this by selecting the semantemes available for this ConcR in conformity with the rules of **L** and uniting them in an appropriate way.
- For a given SemR, the Speaker constructs a corresponding deep-syntactic representation [DSyntR] of *E*, selecting the lexemes which the lexicon and the grammar of **L** provide for this SemR and uniting them in an appropriate way.

I will abstract away from the closer-to-surface linguistic operations leading from the DSyntR of *E* to *E* itself.

☛ In this paper, ConcRs and SemRs are presented informally: they are written in English. A ConcR is printed in the `Courier` font and put between angle quotes (« »); a SemR is printed in regular font and put between semantic quotes (‘ ’). The underscoring in the ConcR and SemR shows the communicatively dominant element ‘**e**’: the whole Conc- or Sem-structure can be reduced to ‘**e**’ with loss but without deformation of the informational content. Thus, ‘tall boy’ ⇔ *tall boy,* while ‘tall boy’ ⇔ *The boy is tall* or *the boy’s tallness.*

For an illustration, see Fig.1: the representations of the free expression *E* = *Water boiling point is 100 degrees Celsius* on the three levels, conceptual, semantic and deep-syntactic.

1) **ConcR(*E*)** «water←**1**–boil←**1**–temperature–**2**→100°C»

2) **SemR(*E*)** ‘water←**1**–boil←**1**–point←**1**–equal–**2**→100°C’

3) **DSyntR(*E*)**

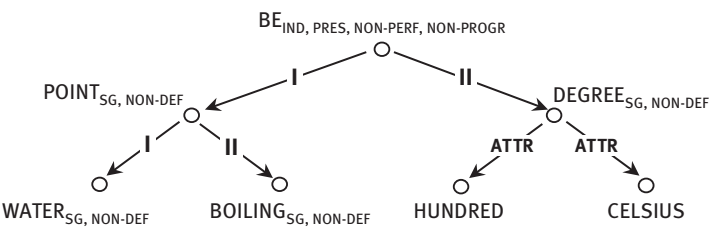


Figure 1: Three levels of representation of a free expression
For technical details, see [Mel’čuk 2012a]; suffice it here to indicate that Roman numbers stand for deep-syntactic actants, and **ATTR** means “deep-syntactic modifier.”

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

Schematically, then, the production of a linguistic expression *E* can be visualized as follows:

(1) ConcR(*E*) ⇐**1**⇒ SemR(*E*) ⇐**2**⇒ DSyntR(*E*)
conceptics **semantics**

To produce a linguistic expression, the Speaker performs two transitions:
1. From prelinguistic concepts to the corresponding linguistic meaning: the Conceptics module.
2. From linguistic meaning to the DSynt-representation of the corresponding expression *E*: the Semantics module.

In the general case, both transitions are performed according to the standard rules of **L** comprising standard conceptual rules of the form «E» ⇔ 'E' and standard semantic rules of the form 'E' ⇔ DSyntR(*E*). For instance:

A standard conceptual rule

«[event X←**1**–]temperature» ⇔ '[X←1–]point**9**'

For instance, as in «[boiling←**1**–]temperature» ⇔ *boiling point*; cf. also *evaporation point*, *freezing point*, *melting point*, etc. The lexicographic number **9** taken from *Longman's Dictionary of Contemporary English Online* identifies this particular sense and the corresponding lexeme POINT**9**.

A standard semantic rule

'[X←**1**–]point**9**' ⇔ [L(X)←**II**–]POINT**9**

☛ Shading indicates the context of a rule, which represents the part of it that is not affected by the rule but is necessary for the rule to apply and to reconnect relevant configurations in the resulting structure.

These rules are called standard because they are the simplest possible. The right-hand part of such a rule contains just one element, namely one semanteme or one lexeme (of course, without counting the elements of the context; the semanteme in the right-hand part of a conceptual rule may appear decomposed). Standard conceptual and semantic rules embody our idealized conception of natural language, according to which a complex linguistic expression is assembled from simple parts that the Speaker selects from storage and puts together—like a car is assembled on a conveyor belt.

The standard conceptual and semantic rules of **L** function in an infinite number of cases, ensuring that both transitions, "ConcR ⇔ SemR" and

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

"SemR ⇔ DSyntR," are mostly free, that is, carried out without individual constraints, concerning particular configurations of semantemes or lexemes. However, natural languages also have a very large number of special cases where these transitions are not free—that is, they cannot be performed following only standard rules. These idiosyncrasies are phrasemes, for which non-standard rules are needed. The number of phrasemes in a language is huge, in the millions, but it is finite. Let us have a closer look at lexical phrasemes—non-free, or constrained, phrases.

1.3 Free phrases *vs*. lexical phrasemes

Definition 1: free phrase

> A phrase *E* of language **L** is free (= non-constrained) if and only if [iff] it is constructed exclusively according to standard conceptual and semantic rules of **L**.

In a free phrase *E*, each of the lexemes L_is is that constitute *E* is selected by the Speaker separately: strictly for its meaning and combinatorial properties and in conformity with the corresponding standard rules of **L** (which are part of **L**'s lexicon, the latter containing also non-standard rules), but without reference to any of the other L_is.

The freedom of a free phrase involves at least two things:

1) The selection of each of L_is not constrained by any other *E*'s lexemic component, so that any lexeme in a free phrase can in principle be replaced by any other semantically and syntactically convenient lexeme without affecting the linguistic correctness of the phrase.
2) The linear order of L_is can be changed according to the general word order rules of **L**.

Definition 2: lexical phraseme

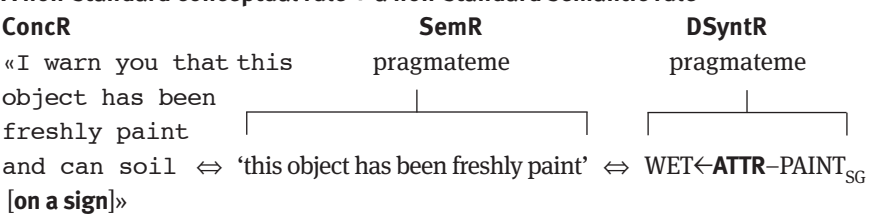
> A phrase E of **L** is a lexical phraseme iff it is not free, that is, iff it is constrained.

The Speaker does not construct a lexical phraseme out of its lexemic components by selecting each one independently of the lexical identity of the others and according only to the standard rules of **L**. In the case of a lexical phraseme he must have recourse to non-standard rules. A non-standard conceptual or semantic rule has as its right-hand part a complex entity—a configuration of semantemes and/or

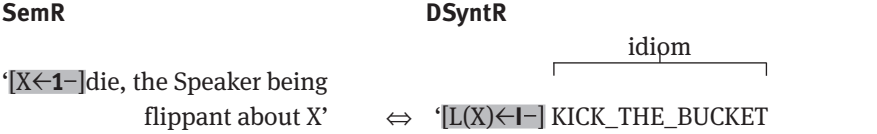
Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

lexemes or a single node to be expanded into a complete surface-syntactic tree. For instance (for pragmatemes and idioms, see below, Sections 3 and 1.5.1):

A non-standard conceptual rule + a non-standard semantic rule



A non-standard semantic rule



A non-standard conceptual/semantic rule specifies a phrase *E* as a whole and marks it (in the SemR and/or the DSyntS) as such in order to block its free manipulation by syntactic and morphological rules, which are not always applicable to constrained expressions. Thus, put simply, a phraseme is a phrase produced either according to a pair "non-standard conceptual rule + non-standard semantic rule" or to a non-standard semantic rule. Non-standard rules normally take precedence over standard rules in a case where both a non-standard and a standard rule are applicable.[1]

1.4 The compositionality of complex linguistic expressions

The discussion and classification of phrasemes requires the notion of compositionality (cf. Mel'čuk 2004).

1 An interesting question, although outside of my topic, is that of conflicts between standard and non-standard rules where both could formally be applied. What happens if the same SemR can be converted to a DSyntR either by a non-standard semantic rule or by a series of standard rules? The result of applying standard rules can be, in some cases, grammatical but not idiomatic. Thus, for the meaning 'X tells Y something that is not true in order to poke fun at Y' the Speaker can say, instead of (i) *X is pulling Y's leg*, something like (ii) *X is poking fun at Y by telling him a lie*. Sentence (ii) is fully understandable and grammatically correct, but sounds un-English. However, it is also possible that a free phrase is as good as a non-free one obtained from the same source. Moreover, under special textual conditions, the same expression can be obtained as a free one and a constrained one.

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

Definitions 3/4: compositional/non-compositional (complex linguistic expression)

These notions are based on the universal operation of linguistic union, symbolized as ⊕. It represents the set of all rules of **L** (its lexicon and its grammar) according to which linguistic units are selected by the Speaker on the paradigmatic axis and then combined on the syntagmatic axis.

> A complex expression $E = L_1—L_2—...—L_n$ of **L** is compositional iff it can be written as $\oplus\{L_i\}$; otherwise, E is non-compositional.

In prose, a complex expression, in particular a phrase, E is compositional iff, according to appropriate rules of **L**, there is a one-to-one correspondence between the components of its signified, those of its signifier and those of its syntactics.[2] In other words, a compositional phrase E can be represented as a regular sum of, respectively, the signifieds, the signifiers and the syntactics of its lexemic components. Among other things, this means that each chunk of E's signified corresponds to a chunk of its signifier (= to a lexeme) and vice versa, as happens in a free phrase, for instance, 'tall$_1$ boy$_2$' ⇔ *tall*$_1$ *boy*$_2$, or in a compositional non-free phrase, such as a collocation (on collocations, see below, Section 1.5.1), for instance,

'coffee$_1$ [without addition of dairy product]$_2$' ⇔ *black*$_2$ *coffee*$_1$.

The meaning 'without addition of dairy product' is carried in this collocation by the adjective BLACK; this is its contextual meaning, that is, a meaning that BLACK has only within this collocation. Note that from the viewpoint of the present discussion, it is irrelevant whether the adjective BLACK receives the sense 'without dairy products' as a separate lexical entry or only in the entry for COFFEE (as a collocate of COFFEE). The only fact that matters is that in the collocation *black coffee* the adjective BLACK is carrying this meaning.

On the other hand, the non-free phrase SEE RED 'be very angry' (an idiom) is not compositional: its meaning cannot be "distributed" between its lexemic components in a natural way following some rules of English.

Compositionality, as follows from Definition 3, is a yes-or-no property, so that a complex linguistic expression cannot be more or less compositional.

Compositionality should not be confused with transparency, which is a psychological and interpretation-oriented, rather than formal and production-oriented, property of phrasemes. Thus, LIGHTNING ROD 'device protecting a structure against

2 Syntactics is the third component of a linguistic sign, namely the set of data on its combinability with other signs, such as the part of speech, the grammatical gender of a noun (in a gender language), the government pattern, and other features of this type; see Mel'čuk 2012a: 47–49.

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

lightning strikes—a conducing object fixed ...' is quite transparent, although non-compositional. Transparency admits of degrees and, for the same expression, can differ for different people. It characterizes a phrase from the viewpoint of its analysis (= understanding). Compositionality, on the other hand, characterizes an expression from the viewpoint of its synthesis (= production). An expression that is fully transparent is necessarily compositional; but the inverse is not true: a compositional expression can be non-transparent. For instance, if you don't know what *yellow press* means, you cannot guess; however, the collocation *yellow press* is compositional: ['press']$_1$ ⇔ *press*$_1$ and '[that distorts the news to create sensations and attract readers']$_2$ ⇔ *yellow*$_2$. The collocation *yellow press* can be represented as a regular sum of the meanings of its lexical components, of their signifiers and of their syntactics.

1.5 General typology of lexical phrasemes

Now I am ready to tackle the typology of lexical phrasemes. To see what types of lexical phrasemes are possible, two situations have to be examined: when Transition 1—"ConcR ⇔ SemR"—is free or when it is not free. (If Transition 1 is free, Transition 2— "SemR ⇔ DSyntR"—must be not free; otherwise, we will have a free expression rather than a phraseme. If Transition 1 is not free, Transition 2 is simply not relevant. In most cases, it is also not free.[3])

1.5.1 Lexemic phrasemes

– Transition 1 (from a ConcR to a SemR) is free, but Transition 2 (from a SemR to a DSyntR) is not free: the linguistic constraints work only for the transition from a meaning to its lexemic implementation. For a given ConcR, the Speaker chooses any meaning 'σ' he wants and the standard rules of **L** allow. But for a selected 'σ', at least one of its possible expressions, *E*('σ'), cannot be constructed according to the standard rules of **L**. For *E*('σ') a non-standard semantic rule is needed. Such a phrase *E* is a lexemic phraseme, since what is constrained (= not free) here is only the construction of the phrase itself (= of its lexemic signifier), not of its meaning.

3 In point of fact, phrasemes for which Transition 1 is not free, while Transition 2 is free do exist. The meaning of such a phraseme is constrained by its referent, but this meaning itself is expressed freely—all its synonymous expressions are possible. Consider, for instance, the signs *Turn off* ⟨*Switch off*, *Close*⟩ *your cell* ⟨*mobile*⟩ (*phone*); note that on signs they do not write ?*No cell phones* or ?*The use of cell phones is not allowed*. (This phraseme is a pragmateme, see Section 3.) However, such phrasemes are not frequent and do not represent a theoretically special case, which allows me to ignore them in this paper.

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

Definition 5: lexemic phraseme

> A lexical phraseme is a lexemic phraseme iff its signifier (= the set of its constituent lexemes and their arrangement) is constrained only with respect to its meaning, that is, to its semantic representation, while its meaning is free with respect to its conceptual representation.

Lexemic phrasemes fall into two major classes: non-compositional and compositional lexemic phrasemes, in other words, idioms and collocations. For more details and explanations concerning lexemic phrasemes, see (Mel'čuk 2012b).

Definition 6: idiom

> A lexemic phraseme is an idiom iff it is non-compositional.

(2) 'die' ⇔ ⌜PASS AWAY⌝ | ⌜KICK THE BUCKET⌝ | ⌜CLOSE [$A_{(poss)}$(N_X)] EYES FOREVER⌝

☛ Idioms are identified in print by raised half-brackets: ⌜ ⌝.

In (2), an idiom that means 'die' (plus nuances of reverence or irreverence) cannot be constructed out of its lexemic components: PASS, AWAY, KICK, BUCKET, etc., according to the rules of **L**, since the meanings of these lexemes are not present in the meaning of 'die':

'die' ≠ 'pass' ⊕ 'away', 'die' ≠ 'kick' ⊕ 'bucket', 'die' ≠ 'close' ⊕ 'one's' ⊕ 'eyes'.

An idiom is selected as a whole, following the corresponding non-standard rule.

To save space, the three major subclasses of idioms—strong idioms, semi-idioms and weak idioms—will not be considered here; see, for instance, (Mel'čuk 2012b and 2015: 305–309).

Definition 7: collocation

> A lexemic phraseme is a collocation iff it is compositional.

(3) 'die' ⇔ MEET/FIND–[$A_{(poss)}$(N_X)]–**II**→$DEATH_{SG,\ NON\text{-}DEF}$
5 000 Romans met/found their death in the battle (not **encountered*/**discovered their death*).

In the collocations described in (3), the noun DEATH, the base of the collocation, is selected freely for the meaning 'die'; but the collocate, the verb MEET or FIND, is selected in a constrained way as a function of DEATH.[4]

4 Here, the verb—MEET or FIND—carries the semanteme 'violently', since this collocation means only a violent death.

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

Again, to avoid sidetracking, I will not discuss different types of collocations: standard, described by standard lexical functions, and non-standard, involving non-standard lexical functions.

1.5.2 Semantic-lexemic phrasemes

– Transition 1 is not free: the constraints work under transition both from concepts to linguistic meanings and from meanings to lexemic expressions. In such a case, the Speaker should not construct **any** meaning 'σ' he wants for the given ConcR «C», even if this meaning is allowed by the standard conceptual rules of **L**, because he risks producing a formally correct, but not idiomatic enough expression. For this particular ConcR «C», the meaning 'σ' to be expressed is prescribed by **L**, namely, by a particular non-standard rule «C» ⇔ 'σ'. This 'σ' can be unique or there can be several possible 'σ_i'; but not anyone of all 'σ' allowed for by standard rules. Most of the time, Transition 2 is also not free. Therefore, the resulting expression *E*('σ') is a semantic-lexemic phraseme, since both the construction of its meaning 'σ' and that of its lexemic expression *E*('σ') are not free.

☛ In the examples of SemRs in (4) the semantic component 'I signal' is used as an abbreviation—to indicate a particular value of the communicative feature of locutionality (Mel'čuk 2001: 242–251): it means that the Speaker resorts to non-descriptive linguistic expressions, which cannot be interrogated or negated, etc.

(4) **ConcR(*E*)** [= «C»] **SemR(**E**)** [= 'σ'] **DSyntR(*E*) and** *E*
(i) (ii) (iii)

«I will now express the same content I have just expressed, but using different words» ⇔ 'I signal that now I will say the same thing, but using different words' ⇔ IN–**II**→WORD$_{\text{PL, NON-DEF}}$
ATTR→OTHER
in other words
⇔ TO–**II**→PUT–**II**→IT
ATTR→DIFFERENTLY
to put it differently

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

French

⇔ AUTREMENT←**ATTR**–DIRE$_{\text{PAST.PART, MASC, SG}}$
autrement dit lit. 'in.other.way said'

Russian

⇔ INAČE←**ATTR**–GOVORIT´$_{\text{IMPF, CONVERB}}$
inače govorja lit. 'differently saying'

The phrasemes *E*s in (4) are a subclass of clichés, formulemes, see Section 2.3 below.

A semantic-lexemic phraseme is constrained with respect to its ConcR, which is (the representation of) its referent; it is also constrained with respect to its SemR. The non-freedom of the phrases under (4iii) means the following. If the Speaker wants ready-made English expressions, then:

First, for the ConcR in (4i) he **must** select the SemR in (4ii), rather than something equivalent (e.g., 'I signal that the following fragment of my speech means the same as the preceding fragment ...'), since otherwise the expressions in (4iii) will not be obtained. This illustrates the constrained character of the ConcR ⇔ SemR passage.

Second, for the SemR in (4ii) he **must** select exactly these expressions rather than anything fairly synonymous, but not idiomatic, like *using some different expressions* or *if I say this in a different way*. This is the constrained character of the SemR ⇔ DSyntR passage.

Definition 8: semantic-lexemic phraseme

> A lexical phraseme is a semantic-lexemic phraseme iff its meaning and its lexemic implementation are both constrained with respect to its referent—that is, to its conceptual representation.

Like lexemic phrasemes, semantic-lexemic phrasemes are of two major types: non-compositional (nominemes) and compositional (clichés).[5]

A non-compositional semantic-lexemic phraseme is a proper name, which has no meaning and is directly linked to its unique referent as an identifying label of the latter: "A proper name [is] a word that answers the purpose of showing what thing it is that we are talking about, but not of telling anything about it" [J.S. Mill, *A System of Logic*, 1.ii.5]. More precisely, a proper name is the name

5 In some of my previous publications on phrasemes, I was erroneously affirming that semantic-lexemic phrasemes cannot be non-compositional: I simply missed complex proper names.

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

of a particular person or group of people, of a particular location or landscape element, of a particular establishment, or of a particular event or period:

(5) Proper name phrasemes (nominemes)

LEO TOLSTOY [a Russian writer and philosopher]
MAO ZEDONG [a Chinese Communist dictator]
ROLLING STONES [an English rock band]
YASNAYA POLYANA [the mansion where the house of Leo Tolstoy was located]
MEDICINE HAT [a city in Canada]
NEW ORLEANS [a city in the US]
NEW HAVEN [a city in the US]
SEVEN PERSONS [a village in Canada]
BLACK SEA
MILKY WAY
POLAR STAR
BIG DIPPER
WHITE HOUSE
ROARING FORTIES
HUNDRED YEARS' WAR
BATTLE OF THE BULGE

Rus. BOL´ŠAJA MEDVEDICA	lit. 'Big She-bear'	[a grouping of stars: Big Dipper]
Rus. VYSADKA V NORMANDII	lit. 'Normandy landing'	[June 6, 1944, Normandy: D-Day]
Rus. VARFOLOMEEVSKAJA NOČ´	lit. 'Bartholomew's Night'	[August 24, 1572, Paris: Saint Bartholomew's Day massacre]
Fr. LA BATAILLE DE LA MOSCOVA	lit. 'the Battle of the Moskova River' ≡	
Rus. BORODINSKOE SRAŽENIE	lit. 'Borodino Battle'	[September 7, 1812 battle between Napoleon's Great Army and the Russian Army]
Rus. STALINGRADSKAJA BITVA	lit. 'Stalingrad Battle'	[bloody fighting in 1942-1943 between the Nazi Army and the Soviet Army in Stalingrad]
Rus. LEDOVOE POBOIŠČE	lit. 'Battle on the Ice'	[April 5, 1242 battle between Livonian knights and Novgorod troops at Lake Peipus]
Rus. TIXIJ OKEAN	lit. 'Pacific ocean'	[Pacific]

These phrasemes can be called nominemes (from *Lat*. *nomen* 'name').

Definition 9: nomineme

> A semantic-lexemic phraseme is a nomineme iff it is non-compositional.

Informally, a nomineme is a phrase that constitutes a complex proper name. As said above, a nomineme *E* does not have a meaning (its signified is empty), but it is directly linked to a given individual referent **r**: **r** $\Leftrightarrow$ *E*.

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

A nomineme is non-compositional since it does not feature a one-to-one correspondence between the components of its signified and the components of its signifier: its signified is empty, it has only a referent.

The following four relevant properties of nominemes should be mentioned:

1. A nomineme may contain a meaningful word or even be fully composed of meaningful words, but this does not change the fact that a nomineme does not have meaning. It may even include a word that denotes the kind to which its referent belongs: Big River is a river, Mount Everest is a mountain, and Atlantic City is a city. Nevertheless, the phrase BIG RIVER simply is, as a whole, the name of a river, and this name is, formally speaking, as meaningless as MISSISSIPI or THAMES. The meaning that a speaker perceives in some naming phrases is what is called inner form, or "semantic etymology."
2. Being meaningless, a nomineme is represented in a SemS not by a semanteme, but by a phonemic/graphemic string, which occupies one node, for instance, '<u>boy</u>←**1**–name–**2**→*John_Enfield*' (*the boy called John Enfield*). This string is a simple place-holder supposed to identify the nomineme itself, which appears in the DSynt-structure also as one node, supplied with all the information necessary for its subsequent processing by syntactic rules, in the first place, with its surface-syntactic tree (just as the idioms are):

DSyntS **SSyntS**

JOHN_ENFILED ⇔ JOHN$_{\text{(N, proper, masc. first.name)}}$–**name-junctive**→ENFIELD$_{\text{(N, proper, last.name)}}$ ⇔ *John Enfield*

Rus. VELIKIE_LUKI ⇔ VELIKIJ$_{\text{(ADJ, proper, anteposed)}}$←**modificative**–LUKI$_{\text{(N, proper, pl)}}$ ⇔ *Velikie Luki*

[a city in Western Russia]

3. A nomineme may have any surface-syntactic form: *"You're gonna make me lonesome"* is the name of a Bob Dylan's song, *"Saving Private Ryan*," that of a Spielberg film, and *Fr. Aux copains d'abord* lit. 'First, to the pals', identifies a restaurant in France (the source of the name being the title of a song by G. Brassens). For a detailed description of complex proper names, see (Bosredon 2011), and for the formal representation of personal names on different levels, as applied to Spanish, (Vincze & Alonso Ramos 2014).
4. A nomineme *Y* is produced from a semantic structure of the form '[<u>X</u>←**1**–name–**2**→]*Y*', that is, '[X, named] *Y*', where 'X' denotes the closest kind of the targeted referent. As a rule, *Y* can be used alone to mean 'X, named *Y*':

'<u>river</u>←**1**–name–**2**→*Mississippi*' ⇔ MISSISSIPPI and *the Mississippi River*
'<u>river</u>←**1**–name–**2**→*Big River*' ⇔ BIG RIVER, but no **the Big River River*

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

The ungrammaticality of the latter is explainable by an obligatory lexical haplology affecting the two contiguous occurrences of RIVER. In the case of *the city of...* we see, of course, no haplology, because there are no contiguous occurrences of CITY:

'city←**1**–name–**2**→*Boston*' ⇔ BOSTON, and *the city of Boston*
'city←**1**–name–**2**→*Kansas City*' ⇔ KANSAS CITY, and *the city of Kansas City*.

1.5.3 Overview of the typology of the lexical phrasemes

For the sake of clarity, I will give here a diagram representing a general typology of lexical phrasemes.

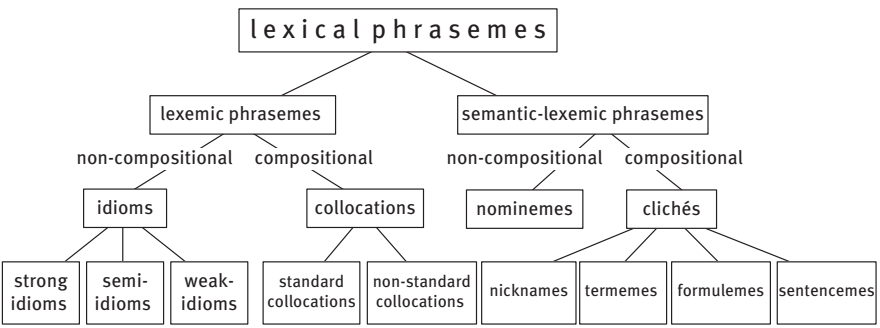


Figure 2: General Typology of Lexical Phrasemes

To facilitate the perception of this classification, let me give examples for the terminal classes of lexical phrasemes.

Lexemic phrasemes
- Idioms
  - Strong: *lame duck* 'elected official who is approaching the end of his tenure'
  - Semi-idioms: *private eye* 'private detective'
  - Weak: *barbed wire* 'artefact used as an obstacle—wire with barbs...'
- Collocations
  - Standard: *meet a criterion* 'be such as the criterion requires'
  - Non-standard: *leap year* 'year that has 366 days'
- Semantic-lexemic phrasemes
- Nominemes *Kansas City*, *Medicine Hat* [a town in Canada], *Charles Dickens*

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

Clichés

Nicknames	*Eternal City* 'Rome', *Big Dipper* 'grouping of 7 stars in the constellation *Ursa Major*'
Termemes	*State Department*, *Foreign Office*, *mad cow disease*
Formulemes	*What time is it?* | *Speedy recovery!*
Sentencemes	*The pen is mightier than the sword.* | *Don't look a gift horse in the mouth.*

It has to be emphasized that the main role of this classification is to establish theoretically clear-cut and formally verifiable divisions between classes of phrasemes. The distribution of phrasemes between these classes is quite a different matter since on many occasions, it is difficult or even impossible to univocally include a given phraseme into this or that class. Such a state of affairs is, however, typical of all our science and technology. Thus, although our measuring units are extremely precise, very often we are unable to properly measure an object, which, however, does not vitiate the measuring units.

I can now consider the class of compositional semantic-lexemic phrasemes—clichés, our main object in this paper.

2 Clichés

Definition 10: cliché

> A semantic-lexemic phraseme is a cliché iff it is compositional.

A cliché has a meaning that is expressed compositionally, that is, this meaning consists of semantic components carried respectively by the cliché's lexemic components. However, both the meaning of a cliché and its lexemic implementation are constrained with respect to the cliché's ConcR, which is the representation of its referent. A classification of clichés can then be based on the four logically possible major types of the constraining referent, defined by two independent features:

the referent is concrete (≈ objective, physical) or abstract (≈ subjective, mental);

the referent is specific (= an individual, in the logical sense of the term) or generic (= a class of individuals).

– If the referent is concrete, it can be:
  1) specific, i.e., an individual, or a particular definite entity of the "outer world," such as a given object or a being, a particular event, a specific substance;
  2) generic, i.e., a class of such individuals.

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

– The referent is abstract; it can be
  3) specific, i.e., an expression of the Speaker's particular wish, intention, attitude, feeling (a ritualized speech act) or a simile concerning a particular situation;
  4) generic, i.e., a general statement by the Speaker about the world (concerning a class of situations).

Thus, we obtain four types of clichés: nicknames, termemes, formulemes, and sentencemes.

Being compositional, a cliché is represented in a DSyntS by as many nodes as it has full lexemes. At the same time, the DSynt-subtree implementing a cliché must be explicitly indicated in order to block the processing of a cliché by regular paraphrasing rules.

2.1 Clichés with a specific concrete referent: proper nicknames

Definition 11: nickname cliché

> A cliché is a nickname cliché iff it has a specific concrete referent (a given individual, in the logical sense).

Nickname clichés, or descriptive names, are similar to nominemes in that they are linked to a single referent; however, they differ from nominemes in that they have meaning: they not only identify their referent but tell us something about it; its components are full, meaningful lexemes. A nickname cliché says exactly what it means but it refers to a single particular person or group of people, a particular substance, a particular location, a particular event, etc. It is a proper nickname.

(6) Nickname clichés

CRAZY HORSE [a famous Lakhota chief, Tasúnke Witkó, 1840–1877, USA]

THE DESERT FOX [Erich Rommel, Nazi general of WWII]

Rus. KORIFEJ VSEX VREMËN I NARODOV lit. 'Coryphaeus of all times and peoples' [Stalin]

Fr. L'HERMITE DE FERNEY lit. 'the hermit of Ferney' [Voltaire]

BROWN PLAGUE [German National Socialism]

RED PLAGUE [International Communism]

ETERNAL CITY [Rome]

GREAT HARLOT [the ancient city of Babylon]

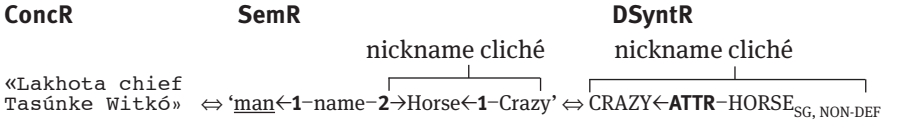
Fr. VILLE (DE) LUMIÈRE lit. 'City (of) Lights' [Paris]

Rus. SEVERNAJA PAL´MIRA lit. 'Northern Palmyra' [Sankt-Petersburg]

Rus. MATER´ GORODOV RUSSKIX lit. 'mother of.cities Russian' [Kiev; with an archaic form **mater´* and non-neutral word order: an adjective following the qualified noun]

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

A nickname cliché, for instance, Crazy Horse, has the following representations:



In the SemR and the DSyntS of this cliché, the lexemes CRAZY and HORSE receive indications of their being part of a cliché in order to restrict their free processing by syntactic rules.

2.2 Clichés with a generic concrete referent: termemes

Definition 12: termeme

> A cliché is a termeme iff it has a generic concrete referent.

A termeme denotes a class of individuals (= of concrete entities or facts). As a typical example, we can take phrases that denote particular diseases: ALZHEIMER'S DISEASE, HUNTINGTON'S DISEASE, CROHN'S DISEASE, etc.

(7)

ALZHEIMER'S DISEASE	'disease of the brain in old people, causing the loss of memory and other mental functions and eventually the death'
HUNTINGTON'S DISEASE	'disease of the brain in people, causing the loss of muscular coordination'
CROHN'S DISEASE	'disease of the bowels in people, causing their inflammation'
LOU GEHRIG'S DISEASE	'disease of the motor muscles in people, causing their inability to function and eventually the death'
LYME DISEASE	'disease in people that, being tick-borne, causes fever, headache, and fatigue'

The bipartite structure of these phrases is obvious: 'disease' is the generic component 'σ_1' of their meaning 'σ' and it is verbalized as DISEASE, while the rest 'σ_2' = '$\sigma - \sigma_1$' of the meaning—the specific differences—is expressed by the name of the discoverer, the sufferer or the place where the disease was identified. Therefore, such phrases could be easily taken to be collocations of the noun DISEASE describable by means of non-standard lexical functions, which specify the collocate. However, this is not the case. In a collocation, the base is selected and used by the Speaker separately, just for its meaning, independently of the collocate; the collocate is added to the base as a function of the latter, when needed, in order to qualify or characterize it. But in the case of

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

ALZHEIMER'S DISEASE etc. the phrase is selected for a given ConcR as a whole to refer to a particular disease. Here, L('σ_2') does not qualify or characterize a disease but **identifies** a particular one. ALZHEIMER'S DISEASE is a single nomination, semantically similar to one-lexeme disease names such as PNEUMONIA, MENINGITIS or JAUNDICE. In contrast, such a phrase as INFECTIOUS/NONINFECTIOUS DISEASE, GASTROINTESTINAL DISEASE or VENEREAL DISEASE is a *bona fide* collocation, where the modifier indicates a subclass of diseases, not a particular disease.

To make my idea clearer, let me present this type of phrase on three levels:

ConcR		**SemR**		**DSyntR**
				termeme
«neurodegenerative disease of the brain that causes the loss of mental functions and death»	⇔	'disease of the brain in old people that causes the loss of memory and other mental functions and eventually the death'	⇔	DISEASE–**ATTR**→ALZHEIMER'S

Termemes are a kind of "natural language terms," linguistic expressions that correspond to the genuine terms in specialized (= technical) languages. Examples of termemes are found mostly among natural species names, anatomical formations, social and political bodies, commercial products, etc.

(8) Termemes

TEA ROSE
KILLER WHALE
DEPARTMENT OF STATE/
STATE DEPARTMENT
FOREIGN OFFICE
Rus. MINISTERSTVO
INOSTRANNYX DEL
lit. 'Ministry of Foreign Affairs'

Rus. ŽELEZNOE DEREVO
lit. 'iron tree' = 'quebracho tree'

Rus. TOLSTAJA KIŠKA lit. 'large bowel' = 'colon'
Rus. SONNAJA ARTERIJA
lit. 'sleepy artery' = 'carotid'
Rus. BERCOVAJA KOST´ lit. '? bone' = 'tibia'
Rus. ŠČITOVIDNAJA ŽELEZA
lit. 'shield-like gland' = 'thyroid'
Rus. GOLLANDSKIJ SYR lit.
'Dutch cheese' = 'type of hard cheese'
Rus. KOLBASA SEMIPALATINSKAJA
lit. 'sausage of Semipalatinsk' = 'type of smoked sausage'

Anyone of these phrases is an "indivisible name" of the corresponding species, formation, social body or product.

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

The following three properties of termemes should be mentioned:

1. In conformity with its nature, a termeme tends to have one-word equivalents in the same or a different language: THYROID GLAND ≡ TYROID, *Rus*. ŠČITOVIDNAJA ŽELEZA ≡ ŠČITOVIDKA 'thyroid', *Rus*. PRJAMAJA KIŠKA ≡ *Eng*. COLON, etc.
2. The modifier component of a termeme is a pseudo-lexeme. Thus, in the termeme *tea rose* the component *tea* carries the meaning 'garden [rose] whose scent is reminiscent of that of tea'. But the lexicon of English should not include the lexeme TEA [X] 'garden [X] whose scent is reminiscent of that of tea', because this meaning is 100% contextual: *tea* has it only in this termeme. ALZHEIMER 'of the brain in old people that causes the loss of memory and other mental functions and eventually the death' should also not be a full-fledged lexeme. (The same is true of many non-standard collocations: BLACK 'without a dairy product' from *black coffee* is also a pseudo-lexeme.)
3. Termemes are easily confused with idioms that designate specific entities. To distinguish termemes from idioms masquerading as termemes, semantic analysis is vital: while termemes are compositional, idioms are not. Thus, the following phrases are all non-compositional and, consequently, not termemes, but termeme-like idioms:

⌜GERMAN SHEPHERD⌝ 'breed of large-size working dogs'
⌜BLACK BOX⌝ 'device in a plane ...'
⌜PRAIRIE DOG⌝ ≡ ⌜GROUND SQUIRREL⌝ 'small rodent living in prairies'
⌜PHONE BOOK⌝ 'listing of telephone numbers in a locality in a book form'

Rus. ⌜BELOE ZOLOTO⌝ lit. 'white gold' = 'cotton'
Rus. ⌜KRAJNJAJA PLOT´⌝ lit. 'extreme flesh' = 'prepuce'
Rus. ⌜SLONOVAJA KOST´⌝ lit. 'elephant bone' = 'substance composing the tusks of an elephant' = 'ivory'
Rus. ⌜VOSPALENIE LËGKIX⌝ lit. 'lung inflammation' = 'pneumonia'

2.3 Clichés with a specific abstract referent: formulemes

A specific abstract referent, that is, a particular given situation, can be many different things. One possibility is an internal state of the Speaker: a wish, a question, a conviction, a statement; the corresponding cliché is a ritualized speech act expressing this state. Another possibility is a moment of time, a specific state of affairs or a specific event in the universe, i.e., in the atmosphere or on the stock market, etc.

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

Definition 13: formuleme

> A cliché is a formuleme iff it has a specific abstract referent.

Since in a formuleme all its lexemes retain their own meaning, its semantic representation is, in a sense, redundant: a formuleme means exactly what it says. Therefore, there is no need to give a formuleme a semantic description in a dictionary (see Section 4).

(9) Formulemes

a. «I wish you everything good in connection with your birthday» ⇔ 'I signal that I wish you a happy birthday' ⇔ *Happy birthday to you!*

b. «I wish you everything good in connection with Christmas» ⇔ 'I signal that I wish you Merry Christmas' ⇔ *Merry Christmas!*

c. «I wish you everything good in connection with my taking leave from you not too close to the end of the day» ⇔ 'I signal that I wish you a nice day' ⇔ *Have a nice day!*

d. «I wish you everything good in connection with your going to perform on stage» ⇔ 'I signal that I wish you to break a leg' ⇔ *Break a leg!*

e. «I wish you to recover as soon as possible from your illness» ⇔ 'I signal that I wish you speedy recovery' ⇔ *Get well soon!* ⇔ *Speedy recovery!*

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

f.

«I ask you to tell me what time it is» ⇔ 'I signal that I want you to tell me what time it is' ⇔ *What time is it?*
⇔ *Do you have time?*

French

⇔ 'I signal that I want you to tell me what hour it is' ⇔ *Quelle heure est-il?*
⇔ *Avez-vous l'heure?*

g.

«I ask you [a woman] to tell me whether you agree to marry me» ⇔ 'I signal that I want you to tell me whether you agree to marry me' ⇔ *Will you marry me?*

Russian

⇔ 'I signal that I want you to be my wife' ⇔ *Bud´ moej ženoj!*

h.

«It is me who uses the italics in this quotation» ⇔ 'Italics is mine' ⇔ *Italics (is) mine.*

French

⇔ 'It is me who emphasizes' ⇔ *C'est moi qui souligne.*

Formulemes are further subdivided according to the type of speech act they express: wishing formulemes, asking formulemes, informing formulemes, etc. An interesting subtype of formulemes are speech formulas (Cowie 2001): formulemes that bear on the actual speech act, that is, that target the utterance of the Speaker or that of the Addressee.

(10) Speech formulemes

a.

«I am surprised by what you just said» ⇔ 'I believe that you should not say this' ⇔ *You don't say.*

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

b.
«Assuming that you understand the fine nuance in what I just said, I am checking whether you do» ⇔ 'if you know what I mean' ⇔ *If you know what I mean.*

c.
«I mockingly dismiss what you just said» ⇔ 'I believe that you are joking' ⇔ *You must be joking.* ⇔ *You are kidding me !*

d.
«I am addressing both men and women in the audience» ⇔ 'I am addressing all the gentlemen and ladies' ⇔ *Ladies and gentlemen!*

Texts swarm with formulemes: *most of the time*; *to give an example*; *It is time to move on*; *It is generally agreed that ...*; *generally speaking, ...*; *We are looking at ...*; *Simply stated, ...*; *X is a factor in Y*; *X makes X's way to/through ...*; *to a greater or lesser extent*; etc.

A formuleme is not necessarily a concrete lexemic expression: it can be a schema of a family of such expressions. Consider, for instance, designations of moments of time:

(11) a. «3: 50» ⇔ 'ten to four' ⇔ *ten to four*

French
⇔ 'four hours minus ten' ⇔ *quatre heures moins dix*
lit. 'four o'clock minus ten'

Russian
⇔ 'four hours without ten' ⇔ *bez desjati četyre*
lit. 'without ten four'

German
⇔ 'ten before four' ⇔ *zehn vor vier*
lit. 'ten before four'

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

Of course, there is no need to write specific rules for each minute; the solution is a general schema:

b. «NUM_1 (hours): NUM_2 (minutes)» ⇔
⇔ 'NUM_3 to NUM_4' ⇔ *NUM_3 to NUM_4* | $\text{NUM}_3 = 60 - \text{NUM}_2$; $\text{NUM}_4 = \text{NUM}_1 + 1$
French
⇔ 'NUM_4 hours minus NUM_3' ⇔ *NUM_4 heures moins NUM_3*
Russian
⇔ 'NUM_4 hours without NUM_3' ⇔ *bez NUM_3 NUM_4*
German
⇔ 'NUM_3 before NUM_1' ⇔ *NUM_3 vor NUM_1*

2.4 Clichés with a generic abstract referent: sentencemes

A generic abstract referent can be a general statement or judgment about the world, the people, etc. Such statements are known as "sayings," "adages," "proverbs," "maxims," "aphorisms," etc. I will call them sentencemes (from *Lat. sententia* 'saying, maxim'). Sentencemes are like formulemes, in that their constituent lexemes have their "expected" meaning and their SemR is redundant in a dictionary.

Definition 14: sentenceme

> A cliché is a sentenceme iff it has a generic abstract referent (a class of situations).

(12) Sentencemes

a. «A person is not completely restricted in his behavior in the presence of a superior» ⇔ 'A cat may look at a king' ⇔ *A cat may look at a king.*

b. «People prefer to be together with people who are similar to them» ⇔ 'Birds having the same type of feather flock together' ⇔ *Birds of a feather flock together.*

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

c.			
«Who is asking for a favor while being in a difficult situation must accept what he is offered»	⇔	'Beggars cannot be choosers'	⇔ *Beggars can't be choosers.*
d.			
«Exaggerated threats are as a rule not carried out»	⇔	'There is big thunder, but there is little rain'	⇔ *Big thunder, little rain.*
e.			
«It is better for something good to happen later than needed than not happen at all»	⇔	'Better late than never'	⇔ *Better late than never.*
f.			
«No one should say something negative about a deceased person»	⇔	'Never speak ill of the dead'	⇔ *Never speak ill of the dead.*
		Latin	
	⇔	'Of the dead you should say either something good, or nothing'	⇔ *De mortuis aut bene, aut nihil.*
g.			
«Good things are usually present in threes»	⇔	'All good things come in threes'	⇔ *All good things come in threes.*

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

h.

«Things are usually present in threes»

French

⇔ 'Two things never appear without a third one' ⇔ *Jamais deux sans trois.*

Russian

⇔ 'God likes the Trinity' ⇔ *Bog troicu ljubit.*

Sentencemes are in fact a motley set. As stated above, they include proverbs, sayings, adages, maxims, aphorisms, etc. The question of the internal organization of the class of sentencemes remains open in spite of good discussions in (Kleiber 2000 and 2010). Luckily, this is not relevant to my goals here. What is important, however, is that a sentenceme is necessarily a complete sentence that expresses a statement and is closed in the sense that it does not have variables for actants, does not normally accept modifiers, etc.[6]

Of course, not all constrained expressions that are full sentences are sentencemes; some of those are idioms, which do not express general judgments about the world but represent statements about a given particular state of affairs (see Klein & Lamiroy 2011: 198):

⌜THE DIE IS CAST⌝ 'The event that cannot be changed has taken place'.
⌜THE CAT IS OUT OF THE BAG⌝ 'The secret has been made known'.
⌜THE MOUNTAIN GAVE BIRTH TO A MOUSE⌝ 'Too many efforts have produced a pitiful result'.

Fr. ⌜LES JEUX SONT FAITS ⌝ lit. 'The bets are down'. = 'Everything has been decided finally'.

6 The proposed description of proverbs, which constitute the core of the sentenceme class, differs from another accepted viewpoint (see, for instance, Goddard & Wierzbicka 2014: 185–204), which considers a proverb to be "word-like" in the sense that its meaning is not the sum of the meanings of its constituent words. I do not think that any proverb is non-compositional and, therefore, an idiom. A proverb can have its **linguistic**, i.e., literal, or compositional, meaning and a "**proverbial**, or formulaic, meaning," which is in fact its Conceptual Representation, or its referent. What is constrained in a proverb is, first, the relation of its linguistic meaning to its conceptual referent (in order to describe the given situation reflected in the proverb only this specific meaning should be selected); and second, the relation of its linguistic form to its linguistic meaning (in principle, no free modifications are allowed in a proverb). I thus share the viewpoint of Kleiber (2000 and 2010) and Tamba (2011) on proverbs.

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

Fr. ⌜UN ANGE PASSE⌝ lit. 'An angel passes'. = 'There is an unexpected and embarrassing pause in a conversation'.

Fr. ⌜À VOS SOUHAITS !⌝ lit. 'To your wishes!' = 'I wish you good health in connection with your sneezing'.

Rus. ⌜TIPUN TEBE NA JAZYK!⌝ lit. 'May an abscess appear on your tongue!' = 'What you have just said may bring disaster, so don't you say such things!'

Rus. ⌜SEDINA V BORODU, BES V REBRO⌝ lit. 'Grey hairs in the beard, the devil in the rib'. = 'An old man feels attracted to much younger women'.

And now, a concluding remark. The distribution of concrete expressions as between idioms and clichés is far from obvious, so that in many cases my specific decisions can be questioned. They are based on what I believe about the perception of these expressions by speakers. Thus, I believe that saying to somebody going to give a recital *Break a leg!* a speaker means 'Break a leg!', while understanding that this is a sort of magic formula used euphemistically to wish a success. By contrast, in *The cat is out of the bag* neither a cat nor a bag are actually meant. If I am wrong, the corresponding expressions have to be reassigned but the proposed classification of phrasemes is not affected.

2.5 Lexicographic description of clichés

A cliché, as has been suggested, is compositional. Its meaning and its lexemic signifier are isomorphic in the sense that each chunk of its meaning corresponds to one of its lexemes (and vice versa), and these lexemes are united following standard syntactic rules in conformity with semantic links between their meanings. In spite of this, a cliché is not constructed by the Speaker from its lexemic components, since it is constrained by its referent. It is selected for this referent and used as a whole according to a non-standard conceptual rule of the language + a non-standard semantic rule (possible exceptions are mentioned in Foonote 3, p. 62). However, a cliché is not a lexical unit and, theoretically speaking, should not have a separate lexical entry; neither should it occupy a single node in a DSynt-tree. Why? Let me consider in this connection the four subclasses of clichés one after another.

– A nickname cliché is a kind of proper name which, although it has a meaning, is attached in one-to-one way to a particular individual referent. Nickname clichés do not belong to the language dictionary—in any event, no more than "normal"

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

proper names and nominemes. They can be relegated to a general encyclopedic dictionary or to a dictionary of proper names.

– In a termeme, the modifying component always carries a non-inherent, contextual meaning that does not qualify the referent of the generic component but identifies a unique particular representative of this referent. Therefore, a termeme needs a lexicographic definition as a whole. However, here its linguistic independence ends. True, a termeme may also have a specific government pattern and restricted lexical cooccurrents, but all these are inherited, to a greater or lesser extent, from the termeme's lexical anchor. The lexical anchor of a cliché is the lexeme that can identify it semantically. This lexeme participates in the cliché's ConcR; a cliché may have more than one lexical anchors. For a termeme, the lexical anchor expresses the generic component of its definition and constitutes its syntactic head. Thus, ALZHEIMER'S DISEASE has DISEASE as lexical anchor and is practically covered by the lexicographic entry of DISEASE as far as its government and collocates are concerned. Therefore, a termeme can be conveniently described in the entry of its lexical anchor using a technique that resembles non-standard lexical functions, being in fact a different phenomenon. To specify a termeme, its ConcR **and** its SemR are given under the lexical anchor and what is supplied is the whole termeme rather than a collocate. For instance, the termeme MAD COW DISEASE is described under DISEASE roughly as follows:

DISEASE
...
«bovine spongiform
encephalopathy» ⇔ 'transmissible lethal
disease of cattle causing
strange behavior' ⇔ *mad cow disease*

– Formulemes and sentencemes need no semantic definitions, since they are compositional and all of their components carry their own inherent meanings. Formulemes and sentencemes have no government pattern and no syntagmatic lexical functions either; they can only have synonyms, antonyms and conversives. They only need a description of their referent, that is, their ConcR. In the examples of (12) the semantic definitions of sentencemes are provided, but, as is immediately clear, they are redundant; they simply repeat the sentencemes themselves. Sentencemes, as complete sentences, have no syntactics; therefore, they are not even linguistic signs in the strict sense of the term. It is not worthwhile supplying formulemes and sentencemes with separate lexical entries. A formuleme/sentenceme can be

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

effectively described in the entry of its lexical anchor(s) in the same way as termemes are. Thus, formulemes (9a) and (9b) are to be specified, respectively, under their lexical anchors BIRTHDAY and CHRISTMAS, sentenceme (12c), under FAVOR (with cross-references under DIFFICULT and ACCEPT), and sentenceme (12f), under DECEASED$_{(N)}$. For instance:

DECEASED, noun, invariant
...
«No one should say something negative about a deceased person» : *Never speak ill of the dead* [**proverb**].

It must be emphasized that the above considerations concerning question of whether or not to give such and such a lexemic expression a separate lexical entry should not be taken too seriously. They are valid only for printed "paper" dictionaries, which are slowly, but steadily disappearing. For the dictionary of the future, an electronic lexicon, this question is utterly irrelevant.

3 Pragmatic constraints and pragmatemes

Lexical phrasemes are defined by lexemic or semantic-lexemic constraints bearing on the production of phrases. However, natural languages also have a further type of constraint, applying not only to complex signs, but to all signs, including individual lexemes. This constraint is as follows: any lexemic expression *E* can be constrained by the situation of its use. This is a pragmatic constraint, and *E* is then a pragmatically constrained expression. It represents a new axis of description, orthogonal with respect to the previous one. The lexicographic description of a pragmatically constrained *E* must include an indication of the appropriate situation of its use, for example, **on an official sign, as a military command, in a letter**, etc.

Pragmatically constrained lexemes

DUCK! [**in a shooting situation**]
FIRE! [**a military command**]
HELP! [**in a situation of danger**]
MEN [**on a bathroom door sign**]
PULL/PUSH [**on a door**]
ROGER 'I understand you' [**in a radio communication**]
Sp. ¡DIGA! lit. 'Say!' = 'Hello!' [**answering the phone**]
Sp. ¿SE PUEDE? lit. 'Can [it] itself?' = 'May I go in?' [**asking the permission to enter**]
Rus. POŽAR! lit. 'Fire!' [**asking for help in the situation of a fire**]

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

Pragmatically constrained idioms

⌜HOLD THE LINE!⌝ [**on the phone**]
⌜AT EASE!⌝ [**a military command**]
Fr. ⌜BUREAU DE TABAC⌝ lit. 'bureau of tobacco' = 'tobacco store' [**on a sign**]
Fr. ⌜EN JOUE!⌝ lit. 'In cheek!' = 'Take aim!' [**a military command**]

Pragmatically constrained collocations

Take aim! [**a military command**]
Wet paint [**on a sign**]
Stop thief! [**asking for help in the situation of a theft on the street**]
Fr. Au voleur! lit. 'To the thief!' = 'Stop thief!' [**asking for help in the situation of a theft on the street**]
Fr. Au feu! lit. 'To the fire!' = 'Fire!' [**asking for help in the situation of a fire**]
Fr. Roulez au pas lit. 'Drive at.the walking.pace' = '(Drive) slowly' [**on a sign**]

Pragmatically constrained lexemes, idioms and collocations are not very numerous. Thus, the main Spanish dictionaries contain, according to approximate calculations in Blanco Escoda 2014, around a few hundred of all these, and we may think that this number is roughly valid for all European languages. But what about clichés? I do not think that pragmatically constrained nominemes, nickname clichés and sentencemes are numerous, if they exist at all. However, pragmatically constrained formulemes form an open-ended set, and for this reason, they deserve a separate name. They will be called pragmatemes.

Definition 15: pragmateme

> A formuleme is a pragmateme iff it is pragmatically constrained.

(13) Pragmatemes
I'll pass the phone to Y [**on the phone**]
To you! [**emptying one's glass of alcoholic beverage**]
Best before ... [**on a container of perishable food**; meaning 'This is best before...']
Sincerely yours [**under a letter**]
Cordially yours [**under a letter**]
For rent [**on a sign**]
Only authorized persons to enter this area [**on a sign**]
Drive slow(ly) [**on a sign**]
No parking [**on a sign;** meaning 'We do not allow you to park here']
Do not enter [**on a sign**]
Wrong way [**on a sign**]
Beware of (the) dog [**on a sign**]

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

Interestingly, the term of pragmateme turned out to be problematic. It can be used either broadly to denote any linguistic unit that is pragmatically constrained, or narrowly, just for pragmatically constrained formulemes. Of course, one can choose any interpretation, provided one defines his use of the term and respects the adopted definition. However, it seems that the broad interpretation is terminologically not felicitous. Under it, all pragmatemes are not phrasemes, but a lexeme can be called a pragmateme. This seems clumsy to me. I like *pragmatically constrained lexeme* better that *lexemic pragmateme*. In this paper, the narrow interpretation is proposed.

Acknowledgments

The present paper has been initiated by the remarks, questions, and objections of F. Louis. He has also read the first version of it, discovering several blunders and confusing passages. Several subsequent versions underwent the scrutiny of L. Iordanskaja, whose criticisms allowed me to correct quite a few mistakes and make my point more clearly and forcefully. The prefinal version was criticized and corrected by M. Alonso Ramos, X. Blanco Escoda, J. Milićević, A. Polguère and two anonymous reviewers for *Yearbook of Phraseology*. I highly appreciate the constructive input received from all these people.

References

Beck, David & Mel'čuk, Igor. 2011. Morphological Phrasemes and Totonacan Verbal Morphology. *Linguistics*. 49(1). 175–228.

Blanco, Xavier. 2014. Inventaire lexicographique d'une sous-classe de phrasèmes délaissée: les pragmatèmes. *Cahiers de lexicologie*. n° 104. 135–156.

Bosredon, Bernard. 2011. Dénominations monoréférentielles, figement et signalétique. In Jean-Claude Anscombre & Salah Mejri (eds.), *Le figement linguistique: la parole entrâvée*, Paris: Honoré Champion, 156–169.

Cowie, Anthony. 2001. Speech Formulae in English: Problems of Analysis and Dictionary Treatment. M/S.

Goddard, Clifford & Anna Wierzbicka. 2014. *Words and Meanings. Lexical Semantics across Domains, Languages and Cultures*. Oxford: Oxford University Press, 185–204.

Fléchon, Geneviève, Paolo Frassi & Alain Polguère. 2012. Les pragmatèmes ont-ils un charme indéfinissable ? In Pierluigi Ligas & Paolo Frassi, eds, *Lexiques. Identités. Cultures*, Verona: QuiEdit, 81–104.

Kleiber, Georges. 2000. Sur le sens des proverbes. *Langages*. 34, n° 139. 39–58.

Kleiber, Georges. 2010. Proverbes : transparence et opacité. *Meta*. 55(1). 136–146.

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

Klein, Jean-René & Béatrice Lamiroy. 2011. Routines conversationnelles et figement. In Jean-Claude Anscombre & Salah Mejri (eds.), *Le figement linguistique: la parole entrâvée*, Paris: Honoré Champion, 195–209.

Mel'čuk, Igor. 2001. *Communicative Organization in Natural Language. The Semantic-Communicative Structure of Sentences*. Amsterdam/Philadelphia: John Benjamins.

Mel'čuk, Igor. 2003. Collocations dans le dictionnaire. In Thomas Szende (réd.), *Les écarts culturels dans les dictionnaires bilingues*, Paris: Honoré Champion, 19–64. See also: http://olst.ling.umontreal.ca/pdf/Collocations-Szende.pdf

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. 2012a. *Semantics: From Meaning to Text*. [Vol. 1.] Amsterdam/Philadelphia: John Benjamins.

Mel'čuk, Igor. 2012b. Phraseology in the Language, in the Dictionary, and in the Computer. In Koenraad Kuiper, ed. *The Yearbook of Phraseology, Vol. 3*. New York/Berlin: de Gruyter Mouton. See also: http://olst.ling.umontreal.ca/pdf/yop_2012_0003.final.pdf

Mel'čuk, Igor. 2013. *Semantics: From Meaning to Text*. Vol. 2. Amsterdam/Philadelphia: John Benjamins.

Mel'čuk, Igor. 2015. *Semantics: From Meaning to Text*. Vol. 3. Amsterdam/Philadelphia: John Benjamins.

Tamba, Irène. 2011. Sens figé: idiomes et proverbes. In Jean-Claude Anscombre & Salah Mejri, (eds.), *Le figement linguistique: la parole entrâvée*, Paris: Honoré Champion, 109–126.

Vincze, Orsolya & Margarita Alonso Ramos. 2014. A Proposal for A Multilevel Linguistic Representation of Spanish Personal Names. In Kim Gerdes, Eva Hajičová & Leo Wanner (eds.), *Dependency Linguistics. Recent Advances in Linguistic Theory Using Dependency Structure*, Amsterdam/Philadelphia: John Benjamins, 119–139.

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM

Brought to you by | Universite de Montreal
Authenticated | igor.melcuk@umontreal.ca author's copy
Download Date | 11/8/15 2:00 AM