Theory and Practice of Lexicographic Definition

Igor Mel'čuk, Alain Polguère

Theory and Practice of Lexicographic Definition

Igor Mel'čuk[I] and Alain Polguère[II]

OLST, Université de Montréal[I]
Université de Lorraine, CNRS, ATILF[II]
igor.melcuk@umontreal.ca[I], alain.polguere@univ-lorraine.fr[II]

Abstract

The paper offers a rigorous characterization of the notion of lexicographic definition: a minimal paraphrase formulated in the same language as the word defined and satisfying six lexicographic principles, which are formulated and discussed. The major types of semantic components in a lexicographic definition are identified and described: firstly, central *vs.* peripheral components, semantic-class marking components, presuppositional components, actant-specification components, weak components, and metaphor-marking components; secondly, conjunctive and disjunctive components. Three additional topics are introduced: the roles that the definition plays in a lexical entry (accounting for the semantic, syntactic and lexically restricted cooccurrence of the headword); lexical units whose definition is problematic; lexical connotations and semantic labels. Though the perspective offered on lexicographic definition is theory-oriented – within the framework of Explanatory Combinatorial Lexicology (ECL) –, a strong emphasis is put on the writing of actual definitions, a couple dozen of which is proposed and analyzed.

Keywords: lexical semantics, lexicographic definition, Explanatory Combinatorial Lexicology, lexicographic methodology

1. Introduction

Lexical meanings form the very core of natural language. First, when we communicate linguistically, we talk to convey informational content – and we do this by means of words. Second, natural language plays a central role in the way we apprehend the world and try to make sense out of it; this again, we achieve by using words. Consequently, the description of word meanings is the most fundamental metalinguistic activity, one that surfaces naturally and constantly in daily life, for instance, when interacting with young children or second language learners:

(1)   – Luscious*, what does it mean?*
     – *Well, it depends, but if you talk about food, luscious food is food that has a very good and rather strong taste.*

The answer given in (1) above is nothing but an improvised definition of *luscious* – at least, of one of its senses. Word definitions are firmly associated with dictionaries, but the practice of defining words – the most essential metalinguistic activity – is probably as old as linguistic communication itself and, in all certainty, is far more ancient than lexicography. The definition of words should therefore concern anyone interested in natural languages, from any perspective: lexicology, of course, but also psychology, acquisition, education, natural language processing, etc. As a consequence, linguistic theory should provide a rigorous methodology for defining word meanings if it aims at extensiveness (coverage of both lexical and grammatical knowledge), applicability and usefulness.

The present paper proposes a detailed characterization of the notion of lexicographic definition in the specific formal framework of Explanatory Combinatorial Lexicology; its approach is therefore very distinct from that of more theory-free writings on the same topic of lexicographic definition, such as Benson *et al.* (1986: Ch. 4). However, our goal is not to review the history and diversity of viewpoints on this notion (the relevant literature is enormous), but rather to present a well-specified vision on how to model word meanings in dictionaries or, more generally, in lexicographic models.

The Explanatory Combinatorial Lexicology (ECL) is an integral component of Meaning-Text theory – namely, the component responsible for lexicological aspects of linguistic description (Mel'čuk *et al.* 1995; Mel'čuk 2013: Ch. 11). In conformity with its fundamental postulate, the representation of the meaning of a lexical unit is its lexicographic definition.

We will elaborate on the notion of lexicographic definition by proceeding in three steps: formal characterization of the definition (Section 2); multiple roles of the definition in a lexicographic entry (Section 3); and lexical units problematic with respect to definition (Section 4). We will conclude by a few remarks on the notion of lexical connotation and that of semantic labels (Section 5). Let us emphasize that the present text updates and sharpens several ECL notions introduced in a number of previous publications, in the first place, Mel'čuk (1988), Mel'čuk *et al.* (1995: 72–117), Mel'čuk (2006) and Mel'čuk & Polguère (2016).

Writing conventions. In accordance with the current ECL practice, the following writing conventions are observed.

- Names of lexical units are printed in small capitals, supplied, when necessary, with lexicographic (i.e. sense-distinguishing) numbers; they may also be accompanied with the specification of their part of speech as subscript, when this is relevant: ROOF$_{(N)}$ **I** [*roof of the palace*], ROOF$_{(N)}$ **II** [*find a roof for my business*], *vs.* ROOF$_{(V)}$ [*huts roofed with tin*], …

- Names of vocables – polysemous words or, more precisely, sets of lexical units that have identical signifiers and are semantically related – are printed in small capitals: the vocable ROOF$_{(N)}$, containing the lexical units ROOF$_{(N)}$ **I**, ROOF$_{(N)}$ **II**, … Homonymous vocables are distinguished by numerical superscripts: BAR$_{(N)}$[1], whose basic lexical unit[1] means 'commercial public place where alcoholic and other drinks can be drunk'; BAR$_{(N)}$[2], whose basic lexical unit means 'piece of something with straight sides'; BAR$_{(N)}$[3], with basic lexical unit meaning 'thing that stops someone from doing something'; BAR$_{(N)}$[4] 'fragment of a piece of music'; BAR$_{(N)}$[5] 'atmospheric pressure measure'; and BAR$_{(N)}$[6] 'profession of lawyers'.

- Names of idioms are enclosed in top corners "⌐…¬": ⌐FRENCH FRIES¬, ⌐ON CLOUD NINE¬, ⌐KEEP LEVEL HEAD¬, ⌐LET THE CAT OUT OF THE BAG¬.

- All linguistic expressions are in italics.

- Linguistic signifieds are put in simple quotes: 'upper covering of a building …' as the signified of ROOF$_{(N)}$ **I**.

- Finally, the names of notions particularly important to our discussion are printed, on the first mention, in sans serif font.

---

[1] The basic lexical unit of a vocable is the lexical unit that semantically controls the polysemy structure of that vocable. For instance, SHARK **I** 'fish …' is the basic lexical unit of the vocable SHARK, because the other lexical unit of this vocable, SHARK **II** 'person who takes advantage of others … as if being a shark **I**', is defined in terms of SHARK **I** as its metaphorical derivative.

2. Formal Characterization of Lexicographic Definitions

2.1   A Definition of Lexicographic Definition

We start with defining the notion of lexicographic definition.

> The lexicographic definition of a lexical unit – from now on, simply *definition* – is a paraphrase of this lexical unit 1) done in the same language as the lexical unit defined, 2) based exclusively on the knowledge of the language itself, and 3) formulated in accordance with the six principles stated in Subsections 2.2 and 2.3 below.

When elaborating the ECL definition for a lexical unit, a lexicographer should not aim to describe the entity/the fact of the real world denoted by this unit. The slogan is: "Think of the word, not of the thing." Conceptual ontologies (Hirst 2009) and world knowledge, no matter how closely they are associated with lexical units, are left out. For instance, the lexical unit ⌜FRENCH FRIES⌝ is defined as 'long thin pieces of potato fried in oil' – without mentioning the facts that the dish originated in Paris just before the French Revolution, was massively introduced into the USA in early 1960s, is utterly unhealthy and was (for a short period) renamed *Freedom Fries* during an American anti-French campaign in 2003. An ECL definition carefully avoids reflecting knowledge that could be qualified as encyclopedic or pragmatic. This does not, of course, prevent a definition from establishing connections between the universe of linguistic semantics and that of extralinguistic concepts: the definition of ⌜FRENCH FRIES⌝ tells us that each "fry" is potato, has a particular form (long and thin) and is fried. The formal modeling of the meaning of a lexical unit can and often does give indirect access to some essential properties of the corresponding concept (Read 1982).

The definitional paraphrase must, as we just said, satisfy six main principles: three of them concern the informational content of a definition (2.2), and three its form (2.3). We will first discuss these principles, after which we will present a typology of semantic components of definitions (2.4).

Note that in our approach the definition of a given headword is not an autonomous component of the corresponding lexicographic entry – in the sense that it must closely collaborate with the description of other properties of the headword, in particular its restricted syntactic and lexical cooccurrence properties (see Section 3).

2.2   Lexicographic Principles Concerning the Informational Content of a Definition

2.2.1   Equivalence Principle

The Equivalence principle (also known as the Adequacy principle, Mel'čuk 1988: 30–31) requires that the definition be a semantic equivalence of two linguistic entities:

$$\text{Definiendum} \equiv \text{Definiens}.$$

Each of the two elements of this equivalence is formally constrained.

- The definiendum is the name of the headword, supplied in the case of a predicative lexical unit with variables which indicate semantic actant slots that the headword controls. Thus, what is defined is not simply the verb KISS(V) **I.1**, but the expression *X kisses* **I.1** *Y on Z*. This expression is called the propositional form of the headword. (The notion of predicative lexical unit is introduced immediately below.)

- The definiens – a linguistic paraphrase of the definiendum – must be the most exact possible. In other words, it is an expression synonymous with the definiendum.

It is a common practice to refer to the definiens as *definition*, which is natural since the definiens is the essential part of the definitional equation. We allow ourselves to do the same when the context excludes any misunderstanding.

Writing Conventions.   In this paper, a "Definiendum ≡ Definiens" equivalence is presented in a box, as shown below for the definition of KISS$_{(V)}$ **I.1**, the basic lexical unit of the corresponding vocable.[2] The definiendum is italicized, since the propositional form is a schema of an utterance (= carrier of a logical proposition) semantically analyzed by the definiens. The latter is not included in semantic quotes ('...'), in order to alleviate the exposition.

X kisses$_{(V)}$ **I.1** *Y on Z*   :   X makes an expressive gesture towards Y
      • X's purpose being to express X's positive feeling towards Y
      • this gesture being X pressing X's lips on a place Z on Y's body
      (, sucking punctually and lightly on Z**)**[3]

Figure 1. Definition of KISS$_{(V)}$ **I.1**

The verb KISS$_{(V)}$ **I.1** is a typical representative of what are known as predicative lexical units: that is, units expressing **semantic predicates** and **quasi-predicates**, which are opposed to **semantic names**.

Semantic predicates are **semantemes**[4] that denote facts – actions, states, events, processes, relations, properties, etc. – and, as a result, are "binding," that is, they control actantial slots and bind their actants into a connected structure (Mel'čuk & Polguère 2008; Polguère 2012). All verbal, adjectival and adverbial lexical units express semantic predicates, along with an important number of nominal lexical units (X's ARRIVAL at Y from Z, X's REGRET over Y, EQUALITY between X and Y). As for quasi-predicates, they denote entities – individuals, objects, substances, etc. –, but have a binding meaning as genuine predicates; they are expressed exclusively by nominal lexical units. The semantically binding nature of quasi-predicates is due to the fact that the corresponding entities are essentially implicated in particular situations from which they borrow their participants. These are expressed as semantic actants of quasi-predicate nouns, which belong to several semantic classes, for instance:

- Qualifying nouns, used rather to characterize than to refer: [individual who is an] IDIOT, PIG **II**, ANGEL **II**, ... (as in *You are an idiot, a pig, an angel, ...*)

- Names of agents and patients: DRIVER [of a car], VICTIM [of a crime], TARGET [of an insult], ...

- Names of parents (kinship terms) and people holding social positions: MOTHER/CHILD [of somebody], MINISTER [of Foreign Affairs], PROFESSOR [of physics], ...

- Names of parts: ARM [of somebody], COVER [of a book], EDGE [of a city], ...

- Names of sets: PACK [of wolves], BUNCH [of idiots], PILE [of books], ...

- Names of substances/materials designed for a particular use: [cleaning] LIQUID, [printing] INK, [fire] WOOD, ...

- Names of objects designed for a particular use: BED [in which somebody sleeps], COMPUTER [which somebody uses to process information], TRAIN [by which somebody transports somebody/something from one place to another], ...

---

2   The vocable KISS$_{(V)}$ is polysemous, with at least the three following senses: KISS$_{(V)}$ **I.1** [*He always kisses his daughter before she goes to sleep.*], KISS$_{(V)}$ **I.2** [*They kissed passionately.*], and KISS$_{(V)}$ **II** [*A gentle breeze was kissing the tree branches.*].

3   The parentheses indicate a weak component of a definition, see Subsection 2.4.1.5 below.

4   A semanteme is the signified of a lexical unit of the language under consideration.

From the viewpoint of their actantial structure, quasi-predicates fall into two major classes: entitative quasi-predicates *vs.* relational quasi-predicates.

1) An entitative quasi-predicate denotes a physical entity (living being, object, substance, place, etc.) involved in a situation: for instance, 'part of X' or 'object used by X for Y'. X, the first semantic actant of such quasi-predicate – as in 'X's medication for Y' – can be expressed as its syntactic dependent: *John's*$_X$ *medication for headache*$_Y$.

2) A relational quasi-predicate denotes an entity X standing in a particular relation to another entity Y. What is specific about this type of quasi-predicate is that X, the first semantic actant of a relational quasi-predicate – as in 'X, who is father of Y' – cannot be expressed as its syntactic dependent, but is readily expressible by means of a copula: *John*$_X$ *is Mary's*$_Y$ *father*.

Lexical units that are neither semantic predicates nor quasi-predicates are semantic names. These are expressed exclusively by nouns whose meanings are not binding and which denote entities: RHINOCEROS, [the] SUN, DANDELION, WATER, etc. The definiendum of a semantic name does not contain actantial variables and is thus formally simpler than that of a predicative lexical unit.

Sun	:	the celestial body
		• that is bright and hot
		• that provides Earth with light and heat
		• that appears, moves across the sky and disappears within regular intervals
		• that is seen as a yellow disk with rays

Figure 2. Definition of SUN

We have dwelt on the predicative/non-predicative nature of a lexical unit in the role of headword because this nature directly conditions its free and restricted cooccurrence (Mel'čuk 2015: Ch. 12; Polguère 2016: 38–41, 301–303). A lexicographic account of the free cooccurrence of a headword consists in specifying its meaning and its syntactic class (roughly, its part of speech), while accounting for its restricted cooccurrence requires much further data.

To sum up: the bulk of the lexicon of any language consists of predicative lexical units, so that consistent application of the Equivalence Principle requires, in the first place, the identification of the actantial slots of the headword if it is predicative. (A good methodology is to presuppose that, by default, the lexical unit under analysis is predicative.)

2.2.2 Semantic Decomposition Principle

The Semantic decomposition principle requires that the definition (that is, the definiens) of a lexical unit be its semantic analysis, that is the decomposition of its meaning into simpler meanings. Let us specify the notion of simpler meaning with an example: the meaning 'egg' is simpler than the related meaning 'omelette' because when one says *omelette* one implicitly expresses the meaning 'egg', but not the other way around.

Semantic decomposition principle entails three consequences.

Consequence 1. The definiens cannot be just formulated as a simple synonym of the definiendum. If this is an exact synonym, what we have is not a definition, but a cross-reference. If this is an approximate synonym, the result is a pseudo-definition, which characterizes the headword approximately; such pseudo-definitions should be proscribed.

Consequence 2. Semantic decomposition in the definiens ensures the absence of vicious circles in the system of definitions. For instance, consider the verb PAINT$_{(V)}$ **I.1** [*He painted the fence.*] and the noun PAINT$_{(N)}$ **I** [*She put three layers of paint.*]. One should not define PAINT$_{(V)}$ **I.1** by means of PAINT$_{(N)}$ **I**

(that is, as PAINT(V) **I.1** ≈ 'cover with paint **I**', as many dictionaries do), while, at the same time, defining PAINT(N) **I** by means of PAINT(V) **I.1** (PAINT(N) **I** ≈ 'colored substance designed to be used to paint(V) **I.1**'). In such a case, the researcher has to determine which of the two lexical units is semantically simpler and use it to define the other. In our pair, PAINT(V) **I.1** is simpler, since to paint does not necessarily imply the use of paint; cf. (2), where the juice of boiled beets does not qualify as PAINT(N) **I**.

(2)   *Margaret painted the fence red with the juice of boiled beets.*

The definition of PAINT(V) **I.1** can now be formulated as follows:

X paints(V) **I.1** *Y color Z with W by means of U*	: X causes **2** that Y becomes of color(N) **I.1** Z[5] • by spreading on Y substance W of color(N) **I.1** Z by means of U • W being liquid

Figure 3. Definition of PAINT(V) **I.1** – Preliminary version

Note that, first, we consider here only the basic lexical unit of a polysemous vocable[6] and second, we will return to this definition later, in Subsection 2.3.1, in order to propose another – more accurate – version of it.

Then, the definition of PAINT(N) **I** is readily formulated:

Z-y paint(N) **I** *for X to use on Y*	: product **I**, designed to be used by X for painting(V) **I.1** Y color(N) **I.1** Z • that is of color(N) **I.1** Z • that is liquid

Figure 4. Definition of PAINT(N) **I**

As one can see, neither definition entails a vicious circle and each represents the meaning of the corresponding lexical units adequately: the action of painting(V) **I.1** does not imply the use of paint(N) **I**, and the paint(N) **I** is a product **I** designed to be used for painting(V) **I.1** (for a detailed discussion of the pair PAINT(V) ~ PAINT(N), see Wierzbicka 1992: 167–169).

Consequence 3. Semantic decomposition, illustrated in the above definitions, functions recursively and leads – sooner or later, but inevitably – to semantemes which are indecomposable, that is, indefinable in terms of semantemes of the same language: these are **semantic primitives**, or **semantic primes** (Wierzbicka 1972, 1985 and many of Wierzbicka's subsequent titles). We will return later to the notion of semantic primes – in Subsection 4.2.4, when discussing the problem of lexicographic description of these very special semantemes.

As far as the depth of decomposition in a definition is concerned, see Subsection 2.3.1 below.

2.2.3   Univocity Principle

The **Univocity principle** requires that all the elements of a definition be univocal, which means that each lexical element in a definition must correspond to one semanteme only and that each semanteme is represented by one lexical element only.

---

5   Two semantemes of causation are distinguished (Kahane & Mel'čuk 2006): 'X causes **1** Y' represents non-voluntary causation, where X is the cause of Y [*The drought caused a severe famine.*]; 'X causes **2** Y' represents voluntary causation, where X is the causer of Y [*Alan slyly caused the failure of the negotiations.*].

6   This wordsense must be distinguished from PAINT(V) **I.2** 'represent by painting(N) **I.1** ...' [*She was painting an apple tree in blossom on the left wall.*]. Lexicographic numbers of the lexical units that appear in the definitions of PAINT(V) **I.1** and PAINT(N) **I** are supplied by ourselves.

This principle boils down to the two following conditions on definitional metalanguage: 1) definitions contain no ambiguity; 2) a given meaning is always expressed in the same way in all definitions – that is, the synonymy of different elements used in definitions is excluded.

In order to meet the first condition, a definition is composed of semantemes and not words. Thus, the definition of TABLECLOTH does not use a vague expression 'table', but the well-specified semanteme 'table$_{(N)}$ **I.1**' (piece of furniture – the basic lexical unit of the corresponding vocable), opposed to the semanteme 'table$_{(N)}$ **I.2**' (the group of people sitting around a table$_{(N)}$ **I.1** [*She lifted the pot, and the whole table screamed!*]):

tablecloth for X to use on Y	:	cover$_{(N)}$ designed to be used by X
		• to cover$_{(V)}$ a table$_{(N)}$ **I.1** Y
		• on which the meal is put
		• that serves to protect Y
		• that is a piece of cloth or cloth-like material

Figure 5. Definition of TABLECLOTH

The condition of non-ambiguity requires to make explicit the designation of each wordsense by using exclusively semantemes, identified by lexicographic numbers. This is not a theoretical problem but rather a practical one. In the description of the whole lexicon of a language – hundreds of thousands of lexical units – a consistent numbering of all wordsenses is impossible without computer tools, which would ensure the management of lexical interconnections within lexicographic entries. (Such an approach is implemented in *Réseau Lexical du Français*, or French Lexical Network, see Polguère 2014.) In this paper, our own definitions are wanting because all elements of their definiens are not genuine semantemes. Thus, in the above definition, the intended wordsenses of COVER$_{(N)}$, DESIGNED, etc. are not properly identified.

The second condition is much harder to satisfy. The lexicographer has to make sure that a given meaning is expressed always – that is, in all definitions – by the same configuration of semantemes. First the lexicographer must identify the semantic content in question and then attach to it a configuration of semantemes that would be acceptable in all possible cases. Let us illustrate with four definitions borrowed from *Longman Dictionary of Contemporary English* (LDOCE 2017):

FILE$_{(N)}$**¹ 4**   a metal tool with a rough surface that you rub on something to make it smooth

PLIERS   a small tool made of two crossed pieces of metal, used to hold small things or to bend and cut wire

SAW$_{(N)}$**² 1**   a tool that you use for cutting wood. It has a flat blade with an edge cut into many V shapes

SCREWDRIVER **1**   a tool with a narrow blade at one end that you use for turning screws

The Univocity principle is violated in these definitions at least in two ways:

- All of the tools *are used* for something, but this fact is not conveyed in the same way in the above definitions: the verb [to] USE is not present at all in the definition of FILE$_{(N)}$**¹ 4**, and in the three other definitions it appears under two forms (*used* or *that you use*).

- The tool parts are not presented in the same way: for instance, in SAW$_{(N)}$**² 1** we have 'It has a [flat] blade', while in SCREWDRIVER **1**, the same meaning is expressed as 'with a [narrow] blade'.

The four lexical units are semantically quite close: they express the same type of semanteme – an entitative quasi-predicate denoting an artifact designed for a particular use (cf. Subsection 2.2.1

above). Therefore, their definitions are supposed to reflect this semantic closeness: they must feature the same generic component (see Subsection 2.3.2 below) and the same general organization. To make a long story short, there is a schematic template for the definitions of tools:

| <L> *for X to use on Y* | : tool I for X to use on Y
• <purpose of use: in order to ...>
• <mode of use: by V-ing ...>
• <structure of the tool I: that consists of ...> |

Figure 6. Definition template for lexical units denoting tools

Here is the definition of FILE(N)[1] 4 written in conformity with this template:

| *file*(N)[1] **4** *for X to use on Y* | : tool I for X to use on Y
• in order to smoothen Y's surface or reshape Y
• by rubbing Y's surface with this tool I in order to remove small amounts of material from it
• that consists of a working part, which is hard and rough, and of a handle |

Figure 7. Definition of FILE(N)[1] **4**

The above definition template allows for the standardization of dozens of definitions since it applies to all lexical units that have the noun TOOL I as their generic term.

The same problem of coherence arises for thousands of definitions: names of inhabitants, names of professions, of emotions, of illnesses, of physical contacts, etc.; verbs of physical state, of communication, of interpersonal relations, of perception, etc. This enumeration shows the path to follow in order to obtain a systematic standardization of semantic configurations: definitions must be elaborated according to semantic classes of the headwords, based on schematic definition template for each of these classes.[7]

2.3 Lexicographic Principles Concerning the Form of a Definition

Having presented the three fundamental principles that constrain definitions from the viewpoint of their content, we switch now to the three principles dealing with the form itself of definitions: they specify the way the definitions must be structured. In actual practice, of course, both aspects – the content and the form of a definition – are inseparable. We have already introduced formal considerations when discussing the definiendum and the definiens; we have even proposed a sketch of the definition template for quasi-predicative nouns denoting tools.

2.3.1 Minimal Decomposition [= Maximal Block] Principle

Each semanteme that is not a semantic primitive is decomposable; therefore, we face the obvious question: to what degree the meaning of the headword must be decomposed in a definition? The ECL prescribes the minimal degree of decomposition – for the following two reasons.

• Maximal decomposition – up to semantic primitives – produces too complex a definiens that requires an excessive effort to be interpreted and understood.

---

7 The requirement of systematic lexicographic processing by semantic classes of lexical units applies not only to definitions; it concerns the lexicographic entries taken globally. Cf. Ju. Apresjan (2008: 57): "To sum up, the third principle of systematic lexicography is the requirement that all salient lexical classes should be fully taken into account and uniformly described in a dictionary in all of their linguistically relevant properties."

- The more one decomposes a definiens, the less the resulting definition can be manipulated as a real paraphrase of the headword.

As a consequence, if we do not want to accept variable decomposition depths (determined by some ad hoc considerations), then the level of semantic decomposition in the definiens has to be minimal. In other words:

> A configuration of semantemes inside a definition which can be expressed by a lexical unit of the language under description – that is, which is lexicalized in this language – must be expressed in the decomposition by the semanteme corresponding to this lexical unit.

This principle is known as the **Minimal decomposition principle** or – in the perspective of lexicalizing the definition – the **Maximal (definitional) block principle**.

The adherence to this principle requires us to revise the definition given in Figure 3, Subsection 2.2.2, for the verb PAINT(V) **I.1**. The generic component of this definition – 'X causes **2** that Y becomes of color(N) **I.1** Z' – is lexicalized in English by the verb COLOR(V) **I.2** [*Alex colored the cake with raspberry extract.*].[8] We are thus obliged to reduce the initial definition of PAINT(V) **I.1**, by reformulating it as follows:

X paints(V) **I.1** *Y color Z with W by means of U*	:	X colors(V) **I.2** Y color(N) **I.1** Z • by spreading on Y substance W of color(N) **I.1** Z by means of U • W being liquid

Figure 8. Definition of PAINT(V) **I.1** – Revised version

Some other approaches than ECL, for instance Natural Semantic Metalanguage (Wierzbicka 1972, 1985, 1992, etc.), resort rather to maximal decomposition in definitions. This is justified by the fact that they concentrate on interlingual correspondences and their final goal is to identify semantic universals, using semantic primitives in the definitions.[9]

2.3.2 Hierarchical Structure Principle

The definiens of a definition consists of configurations of semantemes, called **(semantic) components of the definiens**. These components are organized in a hierarchical structure; the definition must explicitly encode this hierarchy of components. This formal constraint is called **Hierarchical structure principle**.

Terminological remark. The term *component* designates configurations of semantemes, while elementary entities that constitute the definiens (individual semantemes, actantial positions, etc.) are its *elements*. A definiens is thus structured in components, the latter being configurations of elements. The elements of a definiens are "lexical units" of this definiens, while the components are "semantic clauses." (A definiens component may, of course, include just one element.)

The hierarchy of components of a definiens is based on the crucial distinction between a central component and a set of peripheral components. This will be examined in more detail in Subsection 2.4.1.1.

2.3.3 Semantic Network Principle

The **Semantic network principle** concerns the formal nature of the definiendum and the definiens. It stipulates that a definition actually is an equivalence between two semantic networks (Polguère

---

8   COLOR(V) **I.1** is a non-agentive verb seen in *Raspberry extract colored the cake faint pink.*
9   For an introduction to definitions in the framework of *Natural Semantic Metalanguage*, see Goddard (2011).

1997, Mel'čuk 2012: Ch. 4): graphs whose nodes represent semantemes of the language and the arcs, the semantic dependencies connecting predicative nodes (predicates and quasi-predicates) to their actants. Thus, the definition of KISS$_{(V)}$ **I.1**, shown in Subsection 2.2.1, is a verbal "translation" of the equivalence in Figure 9 below.
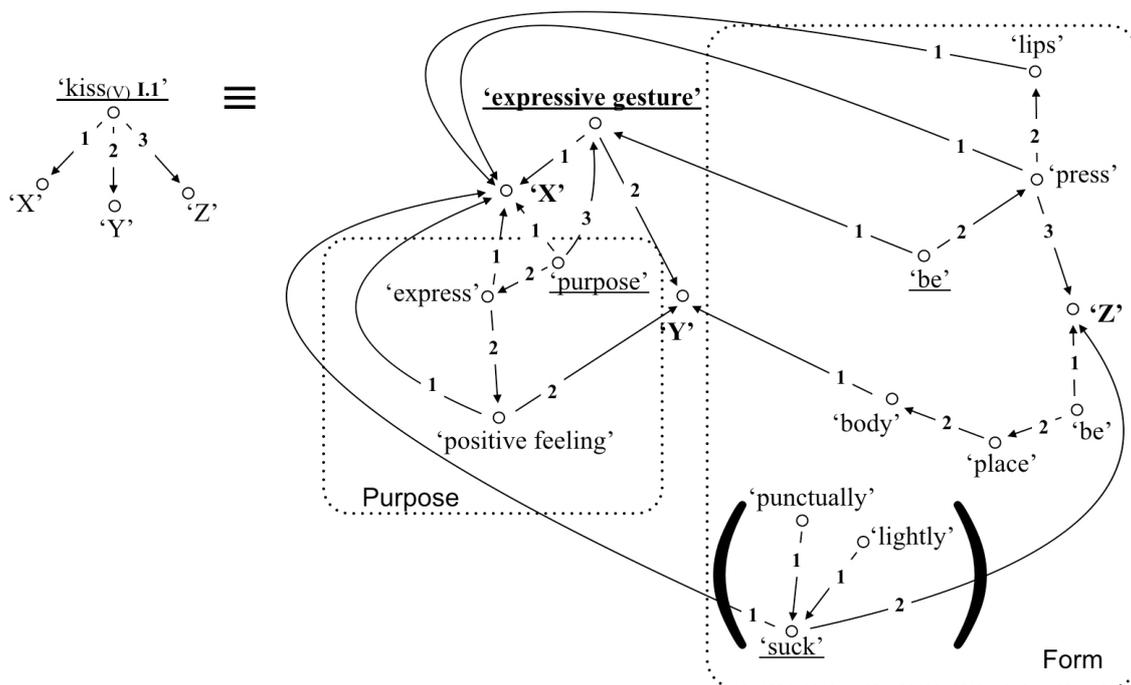

Figure 9. Definition of KISS$_{(V)}$ **I.1** as an equivalence between two semantic networks

The information content of the diagram of Figure 9 corresponds to a genuine definition of KISS$_{(V)}$ **I.1**: it is a complete and precise representation of the meaning of the verb. The right-hand side semantic network in Figure 9 encodes the whole semantic content of KISS$_{(V)}$ **I.1**: all the semantemes, all semantic dependencies, hierarchical structuring of semantic components, their semantic roles ("Purpose" and "Form") and the communicative organization. On this latest point, the underscoring of a node within a given subnetwork shows that the corresponding semanteme is communicatively dominant in the respective semantic substructure (Polguère 1997): thus, the definiens in Figure 9 means, in the first place, '[make] an expressive gesture'. The boldface parentheses in the right bottom quarter of the definiens indicate a weak component, see 2.4.1.5 below.

Beside being a representation of the headword's meaning, the content of Figure 9 is also a formal object from which one can derive many textual formulations, which are equivalent (paraphrastic) definitions of the headword. These include:

- exhaustive definitions – such as those illustrated above, which clearly show the internal structure of the definiens in terms of an enumeration of hierarchized components;

- pedagogical definitions, which should present only essential semantic elements, simplifying or even omitting, if need be, some parts of the definiens.[10]

Moreover, only this type of representation, sufficiently formalized, is accessible for semantic computation, necessary in order to automatically compare lexical meanings, activate polysemy

---

10 On the elaboration of pedagogical definitions in the ECL framework, see Milićević (2016) and Sikora (2016).

patterns for hypothetical wordsense generation à la Pustejovsky (1995)[11] or carry out inferences of the type *He kissed her* ⟹ *He likes her*.

Space limitations prevent us from explaining in more detail the formalization proposed (semantic networks) and from presenting lexicographic tools that allow the researcher to utilize this type of definition in a rigorous and ergonomic way.

The six fundamental principles which underlie the elaboration of definitions in the ECL correspond – more or less – to the problems repeatedly raised in the vast literature on the lexicography of general language dictionaries (Adamska-Sałaciak 2012).

The application of the said principles – in particular, the Hierarchical structure principle – requires a good understanding of the different types of definition components, which leads us to the presentation of a tentative typology of semantic components.

2.4   Types of Semantic Components of a Definition

The semantic components in a definiens can be characterized according to two independent dimensions: the function of the component in the definiens (2.4.1) and its logical structure (2.4.2).

2.4.1   Types of Semantic Components According to their Function in the Definition

2.4.1.1   Central Component *vs.* Peripheral Components

The definiens of a standard definition is divided in two major parts: the obligatory central component and an optional set of peripheral components.[12]

The **central component** is the minimal paraphrase of the headword. There is only one central component per definition and it is obligatory; moreover, its lexicalization has to be of the same deep part of speech as the headword, that is, to have the same passive syntactic valence (see 3.2 below).

The number of **peripheral components** is theoretically unlimited. Each one of them is necessary to express a fragment of the headword's meaning. Taken together, they are sufficient to distinguish the headword by its meaning from all other semantically related lexical units. In our textual definitions, each peripheral component is put on a separate line and preceded by a bullet "•".

Thus, the lexeme CARROT **I.1** [*to peel three carrots*] can be defined as follows, with one central and four peripheral components.

carrot **I.1** *for X* [*to eat*]	: vegetable for X to eat
	• that is sweetish
	• that is orange
	• that is long and pointed
	• that is the root of a plant

Figure 10. Definition of CARROT **I.1**

Each peripheral component is communicatively dominated by the central component or by another peripheral component, which is, in its turn, communicatively dominated by the central component.

---

11 This can be useful for several purposes. For instance, it can be used to help the lexicographer by proposing possible wordsenses for the lexical item under consideration in accordance with the properties of the given semantic class; the lexicographer can confirm or reject these hypotheses. Wordsense generation can also be used in Natural Language Processing to generate a possible interpretation of a word occurrence in a text, to compensate for the fact that all lexicalized senses cannot be expected to be thoroughly described in computerized lexical resources. Furthermore, it can be used in experiments in psycholinguistics to test the plausibility of neological use of words.

12 For some cases of non-standard definitions, see Subsection 4.1 below.

A definition must also encode the semantic nature of the relation between each peripheral component and the component it is subordinated to. Figure 11 presents the hierarchical structure of the definiens in the above definition.

'vegetable for X to eat'			
Taste 'sweetish'	Color 'orange'	Shape 'long and pointed'	Is part of 'root of a plant'

Figure 11. Hierarchical structure of the definiens in the definition of CARROT **I.1**

The distinction "central component *vs.* peripheral components" is, of course, reminiscent of the well-known Aristotelian notion of **definition by genus proximum et differentiæ specificæ** (= 'by the closest kind and specific differences'). We are using here an old recipe, adapting it to definitions of ECL: our central component of a definition roughly corresponds to the generic part of the classical definition, and our peripheral components represent specific differences.

The central component plays a classifying role, allowing for semantic grouping of lexical units that denote facts or entities of a particular class. In the best-case scenario, the central component is composed of just one semanteme, if the language has the corresponding lexical unit – for instance, the central component 'tool **I**' in the definition of FILE$_{(N)}$[1] 4 in Figure 7 (2.2.3 above). It is frequent, however, that such lexicalization is missing, so that it turns out necessary to use a configuration of semantemes as central component in a definiens – for instance, 'expressive gesture' in Figure 9 (2.3.3 above).

Peripheral components, the topic of the next subsections, are of five different types: semantic-class marking components, presuppositional components, weak components, actantial typing components, and metaphor marking components.

2.4.1.2 Semantic-Class Marking Components

The **semantic-class marking component** of a definiens stands in a privileged relation with the central component: it expresses a chunk of the headword's meaning that is crucial to determine the semantic class of the headword and that the central component fails to express.

Let us consider a revealing example: the definition of the noun RAIN$_{(N)}$ **I.1**.[13]

rain$_{(N)}$ **I.1** *on Y*	: water falling from the sky on Y or **1** this falling
	• in the form of drops

Figure 12. Definition of RAIN$_{(N)}$ **I.1** – Preliminary version

This definition is incomplete: it lacks a very important piece of information about the meaning of RAIN$_{(N)}$ **I.1**, namely, that it denotes – like SNOW$_{(N)}$ **I.1**, FOG, THUNDERSTORM, etc. – a particular type of weather phenomenon. It is thus necessary to complement the above definition with a semantic component that expresses the fact that RAIN$_{(N)}$ **I.1** denotes a weather phenomenon – the semantic-class marking component.

rain$_{(N)}$ **I.1** *on Y*	: water falling from the sky on Y or **1** this falling
	• in the form of drops
	• this falling being a weather phenomenon

Figure 13. Definition of RAIN$_{(N)}$ **I.1** – Revised version

---

13 The central component in the definiens contains a disjunction ('or **1**'); disjunctive components are discussed in 2.4.2 below.

Not every definition must have such a component – in fact, all definitions presented before do not. It is needed only if the central component does not do the job of providing the corresponding semantic classification of the headword. On the other hand, the use of this component is not exceptional. For instance, the verbs of the type of MURDER(V), RAPE(V), MUG(V), etc. require in their definition the semantic-class marking component 'this action being a crime': *X murders Y* means 'X kills Y intentionally, this action being a crime'.

2.4.1.3 Presuppositional Components

A lexical meaning 'L' may include a component '⟦P⟧', known as a **presupposition** of 'L'. '⟦P⟧' has the following property:

> If the Speaker says *not L* – that is, if he negates 'L' –, then the component '⟦P⟧' continues to be affirmed (that is, '⟦P⟧' is not accessible to negation).

For instance, *X authorizes Y to do Z* means 'X⟦, who has an authority over Y,⟧ allows Y to do Z'. A little girl can allow (or not allow) her mother to enter her room; but to say *Lola authorizes her mother to enter* implies that the relations between the child and her parents in this family are not normal: the child is presented as having some authority over her parents. The semantic component 'X has an authority over Y' is not negated if the Speaker says *X does not authorize Y to do Z*. If someone says *Lida does not authorize Igor to drink vodka*, he still affirms that Lida has some authority over Igor, as far as alcohol consumption is concerned.

A presupposition in a definition is necessarily a peripheral component: the central component cannot be presupposed. Following our conventions, a presupposition is positioned at the beginning of the definiens – that is, before the central component (which, as we think, facilitates the perception of the definition). Moreover, it is expressed by means of a special syntactic construction: an absolute participial phrase; graphically, it is included in special brackets ⟦ ... ⟧. Figure 14 illustrates this with the definition of the lexeme AUTHORIZE.

	⟦ • X having an authority over Y, ⟧
X authorizes Y to do Z	: X allows Y to do Z

Figure 14. Definition of AUTHORIZE

X authorizes Y to do Z also presupposes 'Y wants to do Z'. However, this other presupposition is part of the meaning 'X allows Y to do Z' and is explicitly expressed in the definition of ALLOW. As a result, it is implicitly present in the definition of AUTHORIZE.

2.4.1.4 Actant Specification Components

An **actant specification component** expresses a semantic constraint on a particular actantial slot. In other words, it specifies the semantic nature of the corresponding actant. The definition of DRINK(V) below will serve to illustrate this, with a component that specifies Y that X is drinking as being liquid.

X drinks(V) Y	: X ingests Y
	• Y being liquid
	• by putting Y into the mouth and swallowing Y
	(• in order to satisfy the physiological need of water in the body)

Figure 15. Definition of DRINK(V)

An actant specification component, such as 'Y being liquid', behaves just like a presupposition in that it is not accessible to negation; as seen in (3) below, this component resists the negation of the verb:

(3)   *Igor does not drink kvas.*

Sentence (3) does not negate the fact that kvas is a liquid – on the contrary, the reader who does not know what kvas is (= a Russian non-alcoholic fermented beverage) learns from (3) that it is a liquid substance.

In spite of its behavior under negation, an actant specification is not a presupposition: the former indicates the semantic class of the actant, while the latter brings in information about a given state of affairs. Thus, the presupposition 'X having an authority over Y' in the definition of AUTHORIZE (Figure 14, 2.4.1.3) does not restrict what semantemes can instantiate the variables X and Y, while 'Y being liquid' in the definition of DRINK restricts the range of potential values of Y.

An actant specification can be semantically so poor – 'X being a living being/an action/a process ...' – that it does not add significant semantic content to the definiens. In such case, it functions strictly for the typing of variables (in the sense used in logic and computer science).

2.4.1.5   Weak Components

The definiens of a headword may contain optionally expressed components: at the conceptual level – the actual situation the Speaker to talking about –, the corresponding situation elements are not necessary for this situation to be referred to by means of the headword. Such components are called **weak components**; a weak component is shown by parentheses in our definitions. For instance, the component 'in order to satisfy the physiological need of water in the body' in the definition of DRINK$_{(V)}$ above (Figure 15, 2.4.1.4) is weak: one can drink something for pleasure (coffee or alcohol), for medical reasons (cough syrup), etc. Nevertheless, 'X drinks Y', considered in isolation, in the first place implies satisfying X's physiological need of water. Additionally, several English expressions manifest a link between 'drink' and 'satisfy a physiological need of water (= thirst)' (e.g., *drink to one's thirst* and *drink thirstily*).

Another stock example of weak component is found in the definition of the verb TEACH:

| *X teaches Y to Z* (*in W*)   :   X acts to cause **2** that Y is learned by Z |
| • Y being structured knowledge |
| • by X communicating elements of Y to Z |
| **(** • in organization W **)** |

Figure 16. Definition of TEACH

X teaches Y to Z (*in W*) is defined as an activity performed in an organization W: a school, a university, a military academy, and the like; if one says *Lida teaches mathematics*, this implies by default that she teaches at a school or a university. However, there are also private teachers, who teach outside of any organization. The above definition remains valid for them too as the component 'in organization W' is weak: it can be dropped any time the Speaker does not need it.

There are several varieties of weak components in definitions; we cannot present all of them, but we would like to dwell on one interesting type widespread in languages that have grammatical gender (Romance and Slavic languages, in the first place). Thus, in French, the masculine noun ÉTUDIANT$_{(MASC)}$ 'student' is applicable to a man only – a female student is denoted by the feminine noun ÉTUDIANTE$_{(FEM)}$. However, things change in the plural: the expression *des étudiants*$_{(MASC)PL}$ can denote a mixed group (for instance, three female and one male students). On the contrary, the expression *des étudiantes*$_{(FEM)PL}$ can refer only to women. Consequently, the component 'of

masculine sex' in the definition of ÉTUDIANT(MASC) is weak, whereas 'of feminine sex' is not weak in the definition of ÉTUDIANTE(FEM). Note that far from all masculine nouns that denote beings featuring a biological gender "lose their masculinity" in the plural: the phrase *les frères*(MASC)PL 'the brothers' denotes only a set of males,[14] while *les cousins*(MASC)PL 'the cousins' can denote a mixed group as well. Similarly, the masculine *taureaux*(MASC)PL 'bulls' denotes only males, but the masculine *chiens*(MASC)PL 'dogs' can denote a mixed group or even a group of female dogs (like in English).

2.4.1.6   Metaphor Marking Components

The definiens of a metaphorical lexical unit of a polysemous vocable must reflect its semantic link to the lexical unit it is a metaphor of within the vocable. This is done by means of a special **metaphor marking component**. For instance, the noun ARM(N) **II** [*the US marketing arm of a Japanese company*] is defined as follows:

[*X,*] *Y-ing arm*(N) **II** *of* Z	:	[X,] a part of Z that is responsible for Y
		• X being a social group
		• Y being an activity of Z
		• Z being a social group
		• ⌐as if¬ X were an arm(N) **I.1** of Z with which Z were doing Y

Figure 17. Definition of ARM(N) **II**

The metaphor marking component, introduced by the expression '⌐as if¬', makes the **semantic bridge** – i.e. a shared semantic component – between ARM(N) **II** and the basic lexeme of the vocable, ARM(N) **I.1** 'upper limb' explicit. In other words, this component indicates the metaphor underlying the **polysemy relation** between ARM(N) **II** and ARM(N) **I.1**. This component is needed to formally link what are known as metaphorical wordsenses to their "source" lexical unit.

More generally, while describing a polysemous vocable the lexicographer must make sure that all lexical units within this vocable are directly or indirectly linked in their definiens by shared semantic components, that is, by semantic bridges. In the case of a metaphorical lexical unit, the metaphor marking component constitutes the necessary semantic bridge.

2.4.2   Types of Semantic Components According to their Logical Structure

Two logical types of semantic components are considered: conjunctive compound components, which contain a logical conjunction (AND **1**), and disjunctive compound components, which contain a logical inclusive disjunction (OR **1**).

A conjunctive compound component is found, for instance, in the definition of the noun WIDOW **1** [*tax deduction for widows*]:[15]

| [*X,*] *widow* **1** | : | [X,] woman |
| | | • whose husband has died and **1** who has not remarried |

Figure 18. Definition of WIDOW **1**

To illustrate a disjunctive compound component, let's take the noun DREAM(N) **I** [*I can't remember my dream.*]. First thing to mention is that this lexeme is the nominalization of the verb DREAM(V) **I** [*I dreamt of you.*], and therefore its definiens must be based on 'X dreams(V) **I** of Y' (for definitions of

---

14 In Spanish, *los hermanos*(MASC)PL lit. 'the brothers' can mean either 'brothers' or 'brothers and sisters'; however, *las hermanas*(FEM)PL 'the sisters' denotes exclusively women.

15 WIDOW **1** is a monoactantial quasi-predicate that denotes a woman with a given "social status" as regards to marriage. It contrasts with WIDOW **2**, a relational biactantial quasi-predicate '[X,] Y's widow **2**' ≡ '[X,] widow **1**, whose deceased husband was Y': *Vladimir's wife is Sasha's widow.*

derivatives, see Subsection 4.1.1 below). Moreover, this nominalization is particular in that it can be used to denote not only the process of dreaming, but also the things dreamt of, that is the actant Y. This leads to the following definition.

| *X's dream*(N) **I** *of Y* | : process of X dreaming(V) **I** of Y or **1** Y dreamt(V) **I** of. |

Figure 19. Definition of DREAM(N) **I**

This disjunctive central component of the above definiens allows for the three following uses of DREAM(N) **I**, namely, to denote: the process of dreaming (4a), or the situation dreamt of (4b), or even both at the same time (4c). This is related to the fact that 'or **1**' stands for inclusive disjunction (≈ and/or).

(4) a. *They say a dream lasts only a few seconds.*   [process]
    b. *My dream was rather bucolic.*   [Y]
    c. *My bucolic dream was brutally interrupted.*   [process and Y]

In terms of lexical functions (Mel'čuk 1996), the noun DREAM(N) **I** is both an S_0 (a "pure," or structural, nominalization) and an S_2 (an actantial nominalization – the typical name of the second syntactic actant) of the corresponding verb DREAM(V) **I**.

Examples of (4c) type prevent the lexicographer from postulating two separate lexemes: one corresponding to S_0 (process) and another being S_2 (the situation dreamt of). The decision to use a disjunctive definition of one lexeme rather than two lexemes is based on the so-called Green-Apresjan Criterion (Mel'čuk 2013: 330–331), also known as zeugma test (Béjoint 1990: 17–18).

Another example of disjunctive definiens component is found in the definition for RAIN(N) **I.1** presented above (Figure 13, 2.4.1.2): 'water falling from the sky on Y or **1** this falling'. The noun RAIN(N) **I.1** can be used to denote the process of water falling (5a), or the water itself (5b), or both at the same time (5c).

(5) a. *The rain lasted two hours.*   [process]
    b. *The rain is cold.*   [water]
    c. *This very cold rain lasted two hours.*   [process and water]

3. Roles the Definition Plays in a Lexicographic Article

We will now touch on a new topic, which is characteristic of the ECL approach: the place of pride the definition takes within a lexical entry. In our perspective, the definition is a focal point for all properties of the headword and especially of its restricted lexical cooccurrence.

As mentioned at the very beginning of this paper, definitions have numerous potential applications – in semantics, in scientific and technical terminography, in language teaching, in Natural Language Processing, etc. But in the first place, each definition plays several important roles from the viewpoint of the coherence of the headword's article. We consider the following three roles of the definition, namely, accounting for:

• the semantic cooccurrence of the headword (3.1);

• its syntactic cooccurrence (3.2);

• its lexical cooccurrence (3.3).

3.1 Accounting for the Semantic Cooccurrence of the Headword

The headword's definition must, first of all, account for the (in-)compatibility of the corresponding semanteme with other semantemes. This is illustrated by the following definition, accompanied by examples (6a–b).

X rinses$_{(V)}$ *Y with Z*	: X cleans$_{(V)}$ Y with Z • Z being liquid • by pouring Z on Y or dipping Y in Z

Figure 20. Definition of RINSE$_{(V)}$

This definition ensures that sentences such as (6a–b) are identified as semantically anomalous:

(6) a. $^{\#}$*Mark rinsed his hands with sand.*
     b. $^{\#}$*Mark rinsed his hands by carefully rubbing them against each other.*

In other words, definitions must prevent the production of absurdities (and tautologies) in texts.

3.2 Accounting for the Syntactic Cooccurrence of the Headword

By introducing the variables for semantic actantial positions, the definition prepares the ground for syntactic actants (on actants, see Mel'čuk 2015: Ch. 12): it determines their maximal number and some of their properties. Thus, the definition of RINSE$_{(V)}$ above allows us to foresee for it a syntactic object, in all probability a direct object, since the communicatively dominant semanteme is 'clean$_{(V)}$' and the verb CLEAN$_{(V)}$ governs a direct object.

Similarly, if the central component of the definiens of a verb is the semanteme 'person X communicates fact Y to person Z', it is quite probable that this verb controls the same actantial positions and that its semantic actants are expressed by the same syntactic actants as those of the verb COMMUNICATE: 'X' ⇔ subject, Y ⇔ direct object and 'Z' ⇔ indirect object. Such is the case with the verbs ANNOUNCE, DECLARE, EXPRESS or REVEAL.

However, this is no more than a loose prediction: even some quasi-synonyms of the communication verbs above show a different syntactic pattern. Such is, for instance, the verb INFORM: 'person X informs person Y of fact Z'.

(7) a. *Polina communicates the news to Nikolay.*
     b. *Polina informs Nikolay of the news.*

We are talking here about predicting, at least partially, the **active syntactic valence** of the headword – that is, the set of its actant slots. However, as mentioned in 2.4.1.1, **passive syntactic valence** (= the set of constructions in which the headword appears as a syntactic dependent), which is a fundamental parameter of syntactic cooccurrence, can also be predicted from the passive valence of the lexical unit expressing the central component of the headword's definition. The passive valence of a lexical unit is determined by its part of speech and its syntactic features; and the headword's part of speech must correspond to the part of speech of the lexical unit expressing the central component of its definiens.

3.3 Accounting for Restricted Lexical Cooccurrence of the Headword

A definition must contain all pieces of information necessary to account for the restricted lexical cooccurrence of the headword. Thus, the verb APPLAUD I [*They applauded him for his brilliant interpretation.*] controls a number of collocations linked to the fact that the intensity and speed of applauding are proportional to the degree of enjoyment/approval of the public:

(8) a. *enthusiastically* <*with enthusiasm*>, *generously*, *heartily*, *loudly*, *rapturously*, *to the echo*, *vigorously*, *warmly*, *wildly*, etc.

   b. *faintly*, *lightly*, *politely*, *reluctantly*, etc.

A large set of collocates expressing the (de-)intensification of the headword's meaning indicates that this headword is prone to (de-)intensification: its definition contains components able to receive this modification (Iordanskaja & Polguère 2005). In the case of APPLAUD **I**, collocations (8a–b) are hooked to the definition components in bold below.[16]

X applauds **I** *Y for Z*	: ⟦ • Y having performed Z in front of X, ⟧ X claps(V) **1** • repeatedly • **to express to Y X's positive feelings toward Z** • **the intensity and speed of X's clapping(V) 1 being proportional to the degree of X's positive feelings**

Figure 21. Definition of APPLAUD **I**

4. Lexical Units Problematic with Respect to Definitions

The definition-writing methodology that has just been presented applies to a vast majority of lexical units. Nevertheless, there are many difficult cases, which are far from being marginal. We will consider in the present section some problematic lexical units, starting with those that are semantically full and proceeding to those that are semantically empty. We will introduce particular descriptive techniques suitable for dealing with such lexical units, which come in two major classes: those which can have definitions (maybe of special type) and those which cannot.

4.1   Lexical Units with Non-Standard Definitions

4.1.1   Structural Semantic Derivatives

A structural semantic derivative[17] of a lexical unit differs from this lexical unit by its part of speech only. Meaning-Text approach postulates five deep parts of speech – V(erb), S(ubstantive), Adj(ective), Adv(erb) and Claus(ative)[18] – appearing in the Deep-Syntactic Structure of sentences. In conformity with this, there are five possible cases of structural derivatives: verbalization, nominalization (≈ substantivization), adjectivalization, adverbialization and clausativization.

The definition of a structural derivative presents us with a paradox:

- if the meaning of a structural derivative is the same as that of its base, its definiens must be the same as that of its base;

- however, the definiens must be of the same part of speech as the definiendum, therefore the definiens of a structural derivative must be different from that of its base.

This paradox is solved in the following way. First, the definiens of the derivative consists of only the central component (without peripheral ones). The core of the definiens is the semanteme corresponding to the base of the derivation, to which we add a redundant semanteme whose

---

16 APPLAUD **II**: *The Committee applauded the Government's efforts*; CLAP(V) **1**: *She clapped twice, and the servant appeared with the meal*; CLAP(V) **2**: *Kids were cheering and clapping*.

17 Semantic derivation does not take into account relations between signifiers, contrary to morphological derivation – the latter being a particular case of the former (Mel'čuk 2015: 160).

18 A clausative is an expression that can constitute a complete clause – all by itself or with its actants, if any: YES, NO, WOW!, DOWN [with N]!, etc.

lexicalization is of the same part of speech as the derivative under consideration. Such is the case, for example, of the definition of the noun APPLAUSE I, which is the nominalization of the verb APPLAUD I, defined in Figure 21 (3.3).

X's applause I *for Y for Z*	:	action of X applauding I Y for Z

Figure 22. Definition of APPLAUSE I

The semanteme 'action' ensures that the definiens of APPLAUSE I is that of a noun, but it is semantically redundant, the meaning 'applaud$_{(V)}$ I' including 'act/action'. This redundancy does not violate any of the definition construction principles introduced in Section 2 – in particular, the Equivalence principle. Note that the semantic networks (2.3.3) for *X applauds* I *Y for Z* and *X's applause* I *for Y for Z* are identical.[19]

Things are different in the case of adjectivalization. Let us concentrate on the important family of adjectives known as **relational adjectives**: SOLAR from SUN, SCHOLAR$_{(Adj)}$ from SCHOOL, INDUSTRIAL from INDUSTRY, etc. Such adjectives have to be defined in terms of their bases, the way it is done for nominalizations. However, strictly speaking, a relational adjective is not a "pure" structural derivative: it is semantically slightly richer than its base. The semantic addition is brought in by the fact of wanting to use a nominal semanteme – e.g. 'Sun' – to characterize another nominal semanteme – e.g. 'matter', as in *solar matter*. It is formally impossible because 'matter' cannot be the first semantic actant of 'Sun', in the same way 'apple' is first semantic actant of 'red' in *red apple*; 'Sun' is a semantic name and cannot have actants. It is thus necessarily the case that the meaning of the phrase *solar matter* contains an "auxiliary" predicative semanteme 'relating [to]', which gives to the definiens a passive adjectival valence:

solar [X]	:	[X] relating to the Sun

Figure 23. Definition of SOLAR

Unlike the component 'action [of]' in the case of nominal derivatives, the element 'relating [to]' is a genuine addition to the meaning of the source of the semantic derivation. This added meaning, very vague, corresponds to what can be called the meaning of an adjective as a particular part of speech. However, in actual speech, the informational content of the phrase *solar X* is a hundred percent function of the values the variable *X* can receive and of additional content that the Speaker supposes inferable by the Addressee in the context of the given enunciation. For instance, the phase *solar matter* mentioned above is used by the Speaker to convey more information than simply 'matter relating to the Sun'. Hearing or reading this phrase, the Addressee is expected to understand something like 'matter the Sun **is composed of**'. The information content shown in bold does not come for the lexical meaning of the adjective SOLAR: it is inferred by the Addressee from the linguistic context the adjective appears in. Additional contents that can be inferred from contexts of use of SOLAR are numerous and variegated, for instance:

(9) a. *solar photon*: 'photon **originating from** the Sun'
  b. *solar flare*: 'flare **located on** the Sun**'s surface**'
  c. *solar system*: '[planetary] system **controlled by** the Sun'

Relational adjectives can also be used to modify predicative nouns of which they fill an actant slot. In this case, neither the 'relating [to]' semantic component of their definiens is activated, nor any additional content needs to be inferred by the Addressee. For instance, with SOLAR:

---

19 Names of events, states, properties, etc., are dealt with in a similar way as names of actions: 'event of ...', 'state of ...', 'property of ...', etc.

(10) a. *solar mass*: 'mass of the Sun'
      b. *solar energy*: 'energy from the Sun'
      c. *solar observations*: 'observations of [= observing] the Sun'
      d. *solar voyage*: '[imaginary] voyage to the Sun'

To sum up, in spite of their differences, the definitions of lexical units such as APPLAUSE **I** and SOLAR are no more than pointers towards their lexical bases. This type of meaning modeling can be called *structural derivational definitions*.

4.1.2   Collocational Lexical Units

Some lexical units are selected by the Speaker as collocates, within a collocation and under control of the collocation base – e.g. *deep* in *deep sleep* (selected as intensifier of the base *sleep*) or *to draw* in *to draw attention* (selected as support verb of the base *attention*). These lexical units can be dubbed **collocational lexical units**.[20] These often have quite standard meaning and their definition presents no problems. Such is, for instance, the case of many lexical units that are elements of the value of the three **syntagmatic lexical functions** for realization verbs: $\mathtt{Real_i}$, $\mathtt{Fact_i}$ and $\mathtt{Labreal_{ij}}$. The notion of lexical function is indispensable in the present paper. However, we cannot offer here a systematic description of this notion, which is available in numerous publications, for instance, Mel'čuk *et al.* (1995: Subsection 3.5) and Mel'čuk (1996, 2015: 155*ff*). Let us simply illustrate what has just been said with the verb DRIVE$_{(V)}$, a $\mathtt{Real_1}$ of many lexical units denoting vehicles. It is defined as follows:

X drives$_{(V)}$ *Y*	:	X controls the movement of Y
		• Y being a vehicle
		• **(** X moving with Y **)**

Figure 24. Definition of DRIVE$_{(V)}$

Lexical units of the type of DRIVE$_{(V)}$ can be selected and used independently of a base – that is, outside of a collocation, as full-fledged, "free" lexical units. There are, however, other lexical units that are **uniquely collocational** and can be used only as collocates, such as AQUILINE (nose), ARTESIAN (well) or PYRRHIC (victory). Their definitions are also not problematic. Thus, AQUILINE is quite naturally defined as follows:

aquiline [X]	:	[nose X] whose shape is curved
		• like the beak of an eagle

Figure 25. Definition of AQUILINE

But between collocational lexical units like DRIVE$_{(V)}$ – which can be used as free lexical units – and uniquely collocational lexical units there is a large class of collocational lexical units whose definition is not straightforward. A good example is the adjective HEAVY used in the following collocations:

---

20 The notion of collocation on which this discussion is based corresponds strictly to phraseological expressions made up of a lexical element "freely" chosen by the Speaker in order to express a given meaning – the **base** of the collocation – and another lexical element that is chosen by the Speaker "under the control" of the base – the **collocate** of the base (Hausmann 1989; Mel'čuk 2015: 320–323).

(11) a. *heavy accent*
     b. *heavy casualties*
     c. *heavy fog*
     d. *heavy accusations*
     e. *heavy costs*
     f. *heavy traffic*
     g. *heavy applause*
     h. *heavy responsibility*
     etc.

At first glance, it might seem that the adjective HEAVY carries different meanings in the above collocations: 'easily perceptible' in *heavy accent*, 'of very serious misdeeds' in *heavy accusations*, 'very enthusiastic' in *heavy applause*, etc. Well, no: in point of fact, these different meanings come from the base rather than the collocate. In all these collocations the adjective HEAVY (with a particular lexicographic number) expresses the same meaning, very general and abstract: that of intensification. A collocational lexical unit of this type must receive a minimal definition – with a definiens as general as possible; thus:

heavy [X]	:	intense [X]

Figure 26. Definition of HEAVY (collocational sense)

Such a definition – an **underspecified definition** – resembles a simple dictionary synonymy cross-reference (e.g., "See syn. *heavy*"). However, it is a pointer to a lexical function. The semanteme 'intense' is one of the semantic triggers of the lexical function `Magn` (Iordanskaja *et al.* 1991), and the above definition indicates that HEAVY is one of the value elements returned under the application of this lexical function:

`Magn`(*applause*) = *deafening, enthusiastic, frenetic, frenzied, **heavy**, loud, raging, rapturous, terrific, thunderous, tumultuous*

`Magn`(*wind*) = *blusterous, blustery, brisk, gusty, **heavy**, high, powerful, strong*

`Magn`(*traffic*) = *bumper-to-bumper, bustling, congested, dense, **heavy***

The elements of the value of `Magn`(L), as one sees, include collocational lexical units that are semantically rather rich – DEAFENING, ENTHUSIASTIC, BLUSTEROUS, BUMPER-TO-BUMPER – and those that are semantically very poor – HEAVY, HIGH, STRONG.

To round up the subsection, let us mention the case of lexical items that are not collocational *per se*, but can be used as a collocate within a single collocation to express a given meaning: for instance, the adjective BLACK 'of black color' expressing the meaning 'without addition of a dairy product' in *black coffee*. In this collocation the inherent, i.e. lexical, or definitional, meaning of the adjective – 'of black color' – is replaced by a contextual collocational meaning of the non-standard lexical function `Without addition of a dairy product`. The existence of this collocation has no consequence on the lexicographic definition of BLACK; it is only relevant to the lexicographic article for COFFEE, which should contain the following information:

`Without addition of a dairy product`(*coffee*) = *black*

This description is justified by the fact that this is an extremely limited usage – just with a single base, COFFEE.[21] Lexical units such as BLACK, that happen to be collocational only in the

---

21 Should this usage become generalized – so that *black* begins to appear with the names of other drinks or dishes to express the meaning 'without addition of a dairy product', we would be forced into postulating a collocational lexical unit BLACK 'without addition of a dairy product'.

lexicographic description of their base, can be qualified as being **contextually collocational**. They receive no lexical entry of their own.

4.2 Lexical Units Without Definitions

4.2.1 Grammemic Lexical Units

Some lexemes in a language with inflection are **grammemic lexical units**: they serve to express analytically particular grammemes or participate in their expression. For example, English has:

- the auxiliary verb HAVE for compound tenses;

- the auxiliary verb BE for the passive voice and the continuous aspect;

- the auxiliary verb WILL for the future tense;

- the adverbs MORE and MOST for the comparative and superlative degrees of adjectives and adverbs;

- the articles, which express the grammemes of definite and indefinite.

The semantic description of grammemic lexical units is not a genuine definition, but a simple cross-reference to the corresponding grammeme, which is itself semantically characterized in the model of the inflection; for instance, for THE:

the [X]	:	$[X]_{\text{DEFINITE}}$

Figure 27. Semantic description of THE

4.2.2 Asemantic Lexical Units

Some lexical units are used not to express meaning, but to play a purely syntactic role in the sentence – that is, in order to ensure that relevant syntactic constraints are observed: these are **asemantic lexical units**. Here are several examples:

- substitute pronouns: ***He***/***She***/***It***/***They*** *fell*, *the person **who** ...*, *the event **which** ...*

- semantically empty pronouns (meteorological or impersonal): ***It** rains*, ***It** is necessary*

- governed prepositions and conjunctions: *insist **on** Y*; *be preoccupied **with** Y*; *ask Y **for** Z*; *believe **that** Y*

Some of these lexical units – for example, THAT$_{\text{(Conj)}}$ in *believe **that** Y* – are semantically empty in all their usages, so that the problem of their definition does not arise. Some others – for example, ON **I** in *insist **on** Y* – carry no meaning in a governed usage, but have an inherent meaning expressed in other contexts, as in (12a–b):

(12) a. *Vladimir is dancing **on** the table.*
b. *If there is snow **on** the ice, extreme caution is advised.*

The lexicographic definition of the corresponding sense of ON – its basic lexical unit ON **I** – is therefore that of a normal full lexeme:

[*X*] *on* **I** *Y*	:	[X] whose location is Y's surface
		• the position of Y being lower than X's position

Figure 28. Definition of ON **I**

4.2.3 "Fossil" (= Idiom-Bound) Lexical Units

Some lexical units can be used only as components of idioms. In most cases, they belong to a previous stage of the language and are synchronically obsolete as independent units. For this reason, they can be termed **"fossil" lexical unit**. Here are a few examples of such lexical units of English:

- the noun BECK, in the idiom ⌜AT BECK AND CALL⌝;

- the noun HUE, in the idiom ⌜HUE AND CRY⌝;

- the adverb FRO, in ⌜TO AND FRO⌝.

These lexical units have to be considered strictly for formal reasons since they are part of the syntactic structure of the idiom they formally belong to; it is, however, impossible to give them a semantic characterization without carrying out an etymological analysis (folk or genuine etymology). Consequently, their lexicographic entries do not include a semantic zone, but are simple cross-references to the respective idioms.

4.2.4 Lexical Units Expressing Semantic Primes

Finally, we will consider a small number of lexemes whose signified is a **semantic prime** – that is, a meaning non-analyzable in the language under consideration in terms of simpler lexical meanings.

A semantic prime cannot be defined by an analytical paraphrase and must be marked as prime in its lexicographic entry. Instead of a *bona fide* definition, it is supplied with an informal explanation and usage examples, which characterize its semantic content at least approximately. Thus, the semantic description of the lexeme NOT – as in ***not** that long ago* –, which expresses a semantic prime, is as follows:

not [*X*]	:	`Negation of` *X* [a semantic prime]

Figure 29. Semantic description of NOT

Now we do not have at our disposal an established list of semantic primes for English. We believe that such a list must be a result of semantic decomposition of many thousands of English lexical units – it is in this way that we hope to obtain the indecomposable semantemes of English. The important particularity of our approach is that we are looking for the semantic primes of individual languages and not for semantic primes of "human thought," as does the Natural Semantic Metalanguage approach (Wierzbicka 1972: 3, 1996 and many more titles; see a recent list of semantic primes in Goddard 2012: 713). We are not sure that linguistic semantic primes are the same for all human languages, but an educated guess as to what could be good candidates for semantic prime status in English can be made. For instance, the following meanings are, in all probability, semantic primes:

'something'	'[a] set' [in mathematical sense]
'more.than'	'space'
'no/not'	'time'
'and **1**' [= logical conjunction]	'say'
'or **1**' [= logical disjunction]	'this.speech.act'

5. Final Remarks on Lexical Connotations and Semantic Labels

For lack of space, we have not discussed such semantic characteristics of the headword that are, so to speak, on the periphery of its meaning and, as a result, are absent from its definition. We mean lexical connotations (5.1) and semantic labels (5.2).

5.1   Lexical Connotations

Take, for instance, the noun WIND$_{(N)}$. It is used not only to denote a movement of atmospheric air, but also to convey several following ideas:

- that of freedom (*free as the wind*);

- that of rapidity (*swift as the wind*);

- that of changing (*The winds of change swept away corruption and incompetence.*);

- that of disappearing (*go with the wind*; *throw/toss something to the wind*);

- that of inconstancy (*see where the wind is blowing*);

- that of vacuity (*words are but wind* [J. Swift]; *windbag*).

The semantemes corresponding to these meanings – 'freedom', 'rapidity', etc. – are not part of the definition of the noun WIND$_{(N)}$: wind, as a weather phenomenon, has no connection with freedom; there are very light and gentle winds; etc. However, these semantemes are closely associated with WIND$_{(N)}$ in the linguistic conscience of English speakers.

To formally represent these associations, the Explanatory Combinatorial Lexicology proposes the notion of **lexical connotation** (Iordanskaja & Mel'čuk 1984, 2009):

> The semanteme 'σ' is a connotation of the lexical unit L if and only if 1) 'σ' is not part of L's definition and 2) the language has a lexical unit L$_\sigma$ whose signifier contains L's signifier and whose signified contains 'σ'.

There are three possible cases of semantic and formal relationships between L$_\sigma$ and L.

1) L$_\sigma$ and L are lexemes of the same vocable – that is, they are linked by the relation of copolysemy. Such is the case of the vocable PIG$_{(N)}$, which contains two L$_\sigma$ lexical units.

- L = PIG$_{(N)}$ **I**: 'farm animal ...'; connotations: voracity, untidiness and unfriendliness.

- L$_{\sigma1}$ = PIG$_{(N)}$ **II.1**: 'person who is voracious – ⌜as if⌝ he were a pig **I**' [linked to PIG$_{(N)}$ **I** by the connotation of voracity].[22]

- L$_{\sigma2}$ = PIG$_{(N)}$ **II.2**: 'person who is untidy and repulsive – ⌜as if⌝ he were a pig **I**' [linked by the connotation of untidiness].

2) L$_\sigma$ is morphologically derived from L. Such is the case of the verb APE$_{(V)}$, derived by conversion from the noun APE$_{(N)}$.

- L = APE$_{(N)}$: 'humanlike wild animal ...'; connotation: mimicking habit.

- L$_\sigma$ = APE$_{(V)}$: 'mimic people in order to make fun – ⌜as if⌝ behaving like an ape' [linked by the connotation of mimicking habit].

3) L$_\sigma$ is an idiom that formally includes L.

- L = WIND$_{(N)}$: 'moving atmospheric air ...'; connotations: freedom, rapidity ...

---

22   For the use of the metaphor-marking semanteme '⌜as if⌝' in definitions, see Subsection 2.4.1.6 above.

- $L_\sigma = \ulcorner$ AS THE WIND $\urcorner$ 1) in the collocation *free as the wind* [linked by the connotation of freedom], 2) in the collocation *swift as the wind* [linked by the connotation of rapidity].

Lexical connotations and the problem of their delimitation with respect to the definition constitute a complicated subject which deserves a separate study and cannot be developed here.

5.2   Semantic Labels

Definitions, as long as they are "textual" representations of their headwords' meaning, are unsuitable for automatic processing of lexical information. They need to be formally encoded, preferably in a representation that is informationally equivalent to their semantic network structure (Subsection 2.3.3). To our knowledge, there are no large-scale lexical resources (electronic dictionaries or lexical databases) that currently incorporates such computer-tractable definitions, for any natural language. However, some level of computation can still be performed based not on full-fledged definitions, but rather on a coarse semantic classification of lexical units. Three examples of such computation are:

- the inheritance of definition template for lexical units belonging to given semantic classes – see the proposed schematic definition template for tools in Subsection 2.2.3;

- more generally, analogical reasoning on lexical models, where new lexicographic information can be inferred from already existing one, based on the assumption that lexical units belonging to the same semantic class tend to share relational and combinatorial properties;

- the semantic annotation of word occurrences in texts – cf. semantic disambiguation –, where a semantic classification, together with paradigmatic and syntagmatic lexical connections for in lexical resources, can be used for the semantic profiling of lexical units.

To satisfy this need for semantic classification, the construction of systems of semantic labels – one for each language under description – has been introduced in ECL methodology (Polguère 2011). The system of semantic labels used in several ECL's lexicographic projects (e.g. Mel'čuk and Polguère 2007) essentially exploits the classifying character of "closest kinds" by conceptualizing the semantic labels as condensed and normalized formulations of the central component of definitions (2.4.1.1), or of their semantic-class marking component (2.4.1.2), if any. It is important to stress that ECL semantic labels are not postulated *a priori* but are inductively identified in the lexicographic process: a given semantic label is proposed when the need arises to semantically classify a set of lexical units whose lexicographic article is being elaborated.

Semantic labels used in lexicography are usually hierarchically structured to form a taxonomic system. Such is the case of the Corpus Pattern Analysis (CPA) Ontology, mentioned in Hanks (2017: 59). In the case of the ECL approach to semantic labeling, only semantic label **classes** – i.e. sets of semantic labels that are semantic derivatives – are hierarchized. For instance, in the current state of the English system of semantic labels, the label class `DIMENSION` is a subclass of the label class `PHYSICAL_CHARACTERISTIC`.[23] The class `DIMENSION` contains the following semantic labels:

- `dimension` – "basic" semantic label of the class, used for nominal lexical units denoting dimensions, e.g. HEIGHT;

---

23   Names of label classes are written in uppercase letters with the underscore sign replacing word spacing; names of semantic labels are written in lowercase letters. Note that, at the present time, only the French hierarchy of semantic labels is made available as a lexicographic by-product of the work on the French Lexical Network (Ollinger & Polguère 2017). The English hierarchy is in its actual state a "translation" of its French counterpart.

- `something/someone that has a certain dimension` – derived label for noms that can be described, in terms of lexical functions, as values of S_1 of nouns denoting dimensions, e.g. GIANT$_{(N)}$;

- `that has a certain dimension` – derived label for adjectives that can be described as values of A_1 of nouns denoting dimensions, e.g. TALL.

Basic labels – that give their name to the corresponding semantic class – are all nominal because only nominal expressions can be used to structure a taxonomy. The main advantage of distinguishing actual semantic labels from semantic classes is that it allows for a global semantic taxonomy of lexical units where lexical units of different parts of speech but semantically very close – such as in the above illustrations – can be grouped together.

Just as lexical connotations (5.1), semantic labels – their identification, structure and use – are a complex topic whose discussion goes beyond the scope of the present study of definitions.

To conclude, let it be emphasized that the definition as proposed and described here has been actively used in ECL for over 30 years, namely in the compilation of Explanatory Combinatorial Dictionaries of Russian (Mel'čuk & Zholkovsky 1984, and its corrected second edition Mel'čuk & Žolkovskij 2016) and of French (Mel'čuk *et al.* 1984-1988-1992-1999). The same type of approach to the definition was independently developed and applied by Ju. Apresjan and his teams in Moscow; see the most recent result: *Active Dictionary of Russian* (Apresjan 2014). The lexicon of a language includes several hundred thousand lexical units; each one requires a definition, and all of the definitions have to be standardized and tuned up to each other. Even for one language this represents a huge amount of lexicographic work. The definitions proposed in this paper – the list of which is summarized in the Appendix – is therefore only a drop in the ocean for the English language.

Acknowledgments

References

Adamska-Sałaciak, Arleta (2012) Dictionary definitions: Problems and solutions. *Studia Linguistica* **129(4)**, 323–339.

Apresjan, Juri D. (2008) Principles of Systematic Lexicography. In Thierry Fontenelle (ed.): *Practical Lexicography. A Reader*, Oxford: Oxford University Press, 51–60.

Apresjan, Jurij (ed.) (2014) *Aktivnyj slovar´ russkogo jazyka. Tom 1 (A–B); tom 2 (V–G)* [Active Dictionary of Russian, Vol. 1 (A–B); vol. 2 (V–G)]. Moskva: Jazyki slavjanskoj kul´tury.

Béjoint, Henri (1990) Monosemy and the Dictionary. In Tamás Magay and Judit Zigány (eds.): *BudaLEX'88 Proceedings. Papers from the 3rd International EURALEX Congress*, Budapest: Akadémiai Kiadó, 13–26.

Benson, Morton, Benson, Evelyn, Ilson, Robert (1986) *Lexicographic Description of English*. Language Companion Series 14, Amsterdam/Philadelphia: John Benjamins.

Goddard, Cliff (2012) Semantic primes, semantic molecules, semantic templates: Key concepts in the NSM approach to lexical typology. *Linguistics* **50(3)**, 711–743.

Hanks, Patrick (2017) Mechanisms of Meaning. In Ruslan Mitkov (ed.): *Computational and Corpus-Based Phraseology (Second International Conference, Europhras 2017 London, UK, November 13–14, 2017 Proceedings)*, Lecture Notes in Computer Science (LNCS) 10596, Berlin/Heidelberg: Springer, 54–68.

Hausmann, Franz Josef (1989) Le dictionnaire de collocations. In Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand and Ladislav Zgusta (eds): *Wörterbücher: ein internationales Handbuch zur Lexikographie/Dictionaries: An International Encyclopaedia of Lexicography/Dictionnaires: Encyclopédie internationale de lexicographie*, Handbücher zur Sprach- und Kommunikationswissenschaft, Band 5.1, Berlin & New York: De Gruyter, 1010–1019.

Hirst, Graeme (2009) Ontology and the Lexicon. In Steffen Staab and Rudi Studer (eds.): *Handbook on Ontologies*. Second edition, International Handbooks on Information Systems, Berlin/Heidelberg: Springer, 269–292.

Iordanskaja, Lidija, Mel'čuk, Igor (1984) Connotation en sémantique et lexicographie. In Mel'čuk *et al.* (1984), 33–43.

Iordanskaja, Lidija, Mel'čuk, Igor (2009) Connotation (in Linguistic Semantics). In Sebastian Kempgen, Peter Kosta, Tilman Berger and Karl Gutschmidt (eds.): *The Slavic Languages. An International Handbook of their Structure, their History and their Investigation*. Berlin/New York: Mouton de Gruyter, 875–882.

Iordanskaja, Lidija, Polguère, Alain (2005) Hooking Up Syntagmatic Lexical Functions to Lexicographic Definitions. In Jurij Apresjan and Leonid Iomdin (eds.): *Proceedings of the Second International Conference on the Meaning Text Theory (MTT'2005)*, Moscow, 176–186.

Iordanskaja, Lidija, Kittredge, Richard, Polguère, Alain (1991) Lexical Selection and Paraphrase in a Meaning-Text Generation Model. In Cécile Paris, William R. Swartout and William C. Mann (eds.): *Natural Language Generation in Artificial Intelligence and Computational Linguistics*. Dordrecht: Kluwer, 293–312.

Kahane, Sylvain, Mel'čuk, Igor (2006) Les sémantèmes de causation en français. *LINX* **54**, 247–292.

LDOCE (2017) *Longman Dictionary of Contemporary English Online*. Pearson ELF, https://www.ldoceonline.com.

Mel'čuk, Igor (1988) Semantic Description of Lexical Units in an Explanatory Combinatorial Dictionary: Basic Principles and Heuristic Criteria. *International Journal of Lexicography* **1(3)**, 165–188.

Mel'čuk, Igor (1996) Lexical Functions: A Tool for the Description of Lexical Relations in the Lexicon. In Leo Wanner (ed.): *Lexical Functions in Lexicography and Natural Language Processing*. Language Companion Series 31, Amsterdam/Philadelphia: John Benjamins, 37–102.

Mel'čuk, Igor (2006) Explanatory Combinatorial Dictionary. In Giandomenico Sica (ed.): *Open Problems in Linguistics and Lexicography*. Monza: Polimetrica, 225–355.

Mel'čuk, Igor (2012) *Semantics: From meaning to text. Volume 1*. Studies in Language Companion Series 129, Amsterdam/Philadelphia: John Benjamins.

Mel'čuk, Igor (2013) *Semantics: From meaning to text. Volume 2*. Studies in Language Companion Series 135, Amsterdam/Philadelphia: John Benjamins.

Mel'čuk, Igor (2015) *Semantics: From meaning to text. Volume 3*. Studies in Language Companion Series 168, Amsterdam/Philadelphia: John Benjamins.

Mel'čuk, Igor *et al.* (1984-1988-1992-1999) *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques. Volumes I-IV*. Montreal: Les Presses de l'Université de Montréal.

Mel'čuk, Igor, Clas, André, Polguère, Alain (1995) *Introduction à la lexicologie explicative et combinatoire*. Paris/Louvain-la-Neuve: Duculot.

Mel'čuk, Igor, Polguère, Alain (2007) *Lexique actif du français. L'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français*. Champs linguistiques, Brussels: De Boeck & Larcier.

Mel'čuk, Igor, Polguère, Alain (2008) Prédicats et quasi-prédicats sémantiques dans une perspective lexicographique. *Revue de linguistique et de didactique des langues (Lidil)* **37**, 99–114.

Mel'čuk, Igor, Polguère, Alain (2016) La définition lexicographique selon la Lexicologie Explicative et Combinatoire. *Cahiers de lexicologie* **109**, 61–91.

Mel'čuk, Igor, Zholkovsky, Alexander (1984) *Explanatory Combinatorial Dictionary of Modern Russian. Semantico-syntactic Studies of Russian Vocabulary*. Vienna: Wiener Slawistischer Almanach.

Mel'čuk, Igor, Žolkovskij, Aleksandr (2016) *Tolkovo-kombnatornyj slovar´ russkogo jazyka. 2-oe izdanie, ispravlennoe* [Explanatory Combinatorial Dictionary of Russian. 2nd corrected edition]. Moskva: Jazyki slavjanskix kul´tur.

Milićević, Jasmina (2016) La définition lexicographique pédagogique: enjeux et difficultés. *Cahiers de Lexicologie* **109**, 93–115.

Ollinger, Sandrine, Polguère, Alain (2017) *Distribution des Systèmes Lexicaux. Version 1.0*. Technical report, ATILF CNRS, Nancy.

Polguère, Alain (1997) Meaning-Text Semantic Networks as a Formal Language. In Leo Wanner (ed.): *Recent Trends in Meaning-Text Theory*. Language Companion Series 39, Amsterdam/Philadelphia: John Benjamins, 1–24.

Polguère, Alain (2011) Classification sémantique des lexies fondée sur le paraphrasage. *Cahiers de lexicologie* **98**, 197–211.

Polguère, Alain (2012) Propriétés sémantiques et combinatoires des quasi-prédicats sémantiques. *Scolia* **26**, 131–152.

Polguère, Alain (2014) From Writing Dictionaries to Weaving Lexical Networks. *International Journal of Lexicography* **27(4)**, 396–418.

Polguère, Alain (2016) *Lexicologie et sémantique lexicale. Notions fondamentales*. Third edition, Montreal: Les Presses de l'Université de Montréal.

Pustejovsky, James (1995) *The Generative Lexicon*. Cambridge, MA/London: MIT Press.

Read, Allen Walker (1982) The Relation of Definitions to their Contextual Basis. *ETC: A Review of General Semantics* **39(4)**, 318–327.

Sikora, Dorota (2016) Activité(s) de définition dans l'apprentissage d'une langue seconde. *Cahiers de Lexicologie* **109**, 117–143.

Wierzbicka, Anna (1972) *Semantic Primitives*. Linguistische Forschungen 22, Frankfurt: Athenäum.

Wierzbicka, Anna (1985) *Lexicography and Conceptual Analysis*. Ann Arbor, MI: Karoma.

Wierzbicka, Anna (1992) Back to definitions: Cognition, semantics, and lexicography. *Lexicographica* **8**, 146–174.

Wierzbicka, Anna (1996) *Semantics. Primes and Universals*. Oxford/New York: Oxford University Press.

Appendix: Lexical Units Dealt With in the Paper (Listed By Semantic Classes)

Lexical units	Location in paper
Semantic predicates	
APPLAUD **I**	Figure 21, 3.3
APPLAUSE **I**	Figure 22, 4.1.1
AQUILINE	Figure 25, 4.1.2
AUTHORIZE	Figure 14, 2.4.1.3
DREAM(N) **I** [also a quasi predicate – cf. disjunction in the central component]	Figure 19, 2.4.2
DRINK(V)	Figure 15, 2.4.1.4
DRIVE(V)	Figure 24, 4.1.2
HEAVY [collocational]	Figure 26, 4.1.2
KISS(V) **I.1**	Figure 1, 2.2.1; Figure 9 [semantic network], 2.3.3
NOT [semantic prime → non-definitional semantic description]	Figure 29, 4.2.4
ON **I**	Figure 28, 4.2.2
PAINT(V) **I.1**	Figure 8, 2.3.1
RINSE(V)	Figure 20, 3.1
RAIN(N) **I.1** [also a quasi predicate – cf. disjunction in the central component]	Figure 13, 2.4.1.2
SOLAR	Figure 23, 4.1.1
TEACH	Figure 16, 2.4.1.5
Semantic quasi-predicates	
ARM(N) **II**	Figure 17, 2.4.1.6
CARROT **I.1**	Figure 10, 2.4.1.1
FILE(N)^1 **4**	Figure 7, 2.2.3
PAINT(N) **I**	Figure 4, 2.2.2
TABLECLOTH	Figure 5, 2.2.3
WIDOW **1**	Figure 18, 2.4.2
Definition pattern for lexical units denoting tools	Figure 6, 2.2.3
Semantic name	
SUN	Figure 2, 2.2.1
Grammemic lexical unit [→ Non-definitional semantic description]	
THE	Figure 27, 4.2.1

Journal of Cognitive Science **19(4)**, 2018, pp. 417–470.

30/30