




**Tilman Reuther**

Universität Klagenfurt  
Austria

 <https://orcid.org/0000-0002-4475-9014>

# Remarks on Ju. D. Apresjan's lexicographic terminology: *leksema* and *vokabula* from a multilingual perspective

## Abstract

This paper discusses the Russian terms *vokabula* and *leksema*, as well as related concepts introduced to lexicography and semantics by a group of Moscow-based researchers led by Jurij D. Apresjan (commonly known as the Moscow Semantic School) and by Igor Mel'čuk (the author of the Meaning ⇔ Text Theory and, alongside Aleksander Zholkovskij, the primary proponent of the Explanatory-Combinatorial Dictionary). The Russian term *leksema* and its equivalents in other languages (Eng. *lexeme*, Fr. *lexème*, Ger. *Lexem*, Sp. *lexema*) are firmly anchored in Apresjan's and Mel'čuk's lexicographic theory and practice and have been widely accepted by the international lexicographic community. In contrast, the Russian term *vokabula* has not even been accepted by Russian lexicographers, let alone the international lexicographic community. This paper explores the reasons behind this discrepancy, which can be attributed to the nature of the lexical parallels in Russian and other languages (Eng. *vocable*, Fr. *vocable*, Ger. *Vokabel*, Sp. *vocablo*). Related issues in this paper include the widely used Russian term *slovarnaja stat'ja* (Eng. *dictionary entry*, Fr. *article de dictionnaire*, Ger. *Wörterbuchartikel*, Sp. *artículo lexicográfico*) and the Russian term *lemma* (Eng. *lemma*, Fr. *lemme*, Ger. *Lemma*, Sp. *lema*), as well as the German term *Stichwort* (Eng. *keyword*), the former in the context of computational linguistics and the latter in the context of general lexicography. As a result, we propose names for Apresjan's and Mel'čuk's term *vokabula* in German (namely, *Stichwort* or *Lemma*) and English (namely, *keyword* or *lemma*).

## Keywords

linguistic terminology, lexicography, lexical parallels, Russian, English, French, German, Spanish

## Introduction

This paper begins with the definitions of the Russian terms *leksema* and *vokabula*, as well as related terms. These were introduced to Russian lexicography and semantics in the 1970s by a group of Moscow-based researchers led by Jurij D. Apresjan (commonly known as the Moscow Semantic School) and by Igor Melčuk (the author of the Meaning  $\leftrightarrow$  Text Theory and, together with Aleksander Zholkovsky, the main proponent of the Explanatory-Combinatorial Dictionary – see Melčuk & Reuther, 2020). The two terms were not equally accepted by the scientific community. The Russian term *leksema* and its equivalents in other languages (Eng. *lexeme*, Fr. *lexème*, Ger. *Lexem*, Sp. *lexema*) – firmly anchored in Apresjan’s and Melčuk’s lexicographic theory and practice – have been widely accepted by the international lexicographic community, albeit not always with the same clarity of definition or consistency of usage found in the work of Apresjan and Melčuk. They are often used as quasi-synonyms of the term *word*. Not so the Russian term *vokabula*. Anchored in Apresjan’s and Melčuk’s lexicographic terminology, this term has not even been accepted by the wider Russian lexicographic community, let alone lexicographers and linguists worldwide. We will argue that this is due to the nature of the lexical parallels<sup>1</sup> in Russian and other languages (English *vocable*, French *vocable*, German *Vokabel*, Spanish *vocablo*). Data from large corpora of Russian and German, as well as general dictionaries of Russian, German, French, Spanish and English, will be used to support this claim. This paper will also discuss the widely used Russian term *slovarnaja stat’ja* (Eng. *dictionary entry*, Fr. *article de dictionnaire*, Ger. *Wörterbuchartikel*, Sp. *artículo lexicográfico*) and the Russian term *lemma* (Eng. *lemma*, Fr. *lemme*, Ger. *Lemma*, Sp. *lema*), as well as the German term *Stichwort* (Eng. *keyword*), the former in the context of computational linguistics and the latter in the context of general lexicography.

The paper is organized as follows: Section 1 is documentary in nature and includes some commentary. Subsection 1.1 covers Apresjan’s definitions; Subsection 1.2 deals with a recent handbook article on “Lexeme” in the authoritative ESLL (*Encyclopedia of Slavic Languages and Linguistics Online*, see Greenberg & Grenoble, 2020–); Subsection 1.3 covers Melčuk’s definitions. Subsection 1.4 provides an intermediate summary. Section 2 is a corpus- and dictionary-based research

---

<sup>1</sup> The term *lexical parallel* is used as defined by Dubichynskyi & Reuther (2020). It refers to two or more words in different languages that are phonetically and/or graphically similar, and that have either the same, or partly different, or completely different lexical meaning. These are called full, partial, or false lexical parallels, respectively.

into the topic addressed in the title: *leksema* and *vokabula* from a multilingual perspective. Subsection 2.1 compares corpus and dictionary data on Russian *leksema* and *vokabula*, on the one hand, and on German *Lexem* und *Vokabel*, on the other hand; Subsection 2.2 contains general dictionary data on English *vocable*, French *vocable* and Spanish *vocablo*. The short Section 3 is devoted to the related terms English *lemma* and *keyword*, German *Lemma* and *Stichwort*, which are widely used in general, computational and corpus linguistics. This leads to Section 4, in which we propose names for Apresjan's and Mel'čuk's term *vokabula* in German (*Stichwort* or *Lemma*) and English (*keyword* or *lemma*). The paper ends with the Conclusion and the Bibliography.

## Section 1: Russian, French and English terminology

### 1.1 Apresjan (2009; 2014): the Russian terms *leksema* and *vokabula*

Let us begin with Apresjan's definition of the term *leksema* as the core concept of his "Conceptual apparatus of systemic lexicography" (Apresjan, 2009, p. 509; translation mine – T.R.<sup>2</sup>):

"A *leksema* is a word in one of its senses, together with all its inherent properties in this sense, essential properties being those to which the rules of the given language refer. [...] It is the *leksema*, not the word, that is the real lexical unit of the language, since normally every polysemic word is used in an utterance in one of its senses. Unlike the meaning, the *lexeme* is a multifaceted unit of language: it has a signifier (phonetic shell), a signified (meaning) and a syntactics (features of combinability)."<sup>3</sup>

---

<sup>2</sup> For the sake of accuracy and transparency, the relevant Russian terms appear in italics and transliteration, rather than in translation from the original Russian.

<sup>3</sup> The original Russian text reads as follows: «Лексема – слово в одном из его значений, но во всей совокупности присущих ему в этом значении свойств, причем существенными свойствами признаются те, к которым обращаются правила данного языка. [...] Именно лексема, а не слово, является реальной лексической единицей языка, потому что в норме каждое многозначное слово используется в высказывании в каком-то одном из своих значений. В отличие от значения, лексема – многосторонняя единица языка: у нее есть означающее (фонетическая оболочка), означаемое (значение) и синтактика (особенности сочетаемости)».

Apresjan's definition of the *leksema* highlights the traditional structuralist concept of the dual nature of the linguistic sign: the signifier (Rus. *označajuščee*) and the signified (Rus. *označaeomoe*). This concept was enriched by Melčuk's syntactics (Rus. *sintaktika*) as the third component of the linguistic sign in Meaning ⇔ Text Theory (Melčuk, 1974).

This definition of *leksema* relies on the concepts of *word* (Rus. *slovo*) and *meaning* (Rus. *smysl*), as well as the ability to identify the (signified) meaning of one (signifying) element within the (signified) meaning of a (signifying) utterance. Further on, the concept of *polysemy* (Rus. *mnogoznačnosť*) of words is crucial for linguists when compiling dictionaries. Apresjan also mentions his approach to linguistic modelling – the concept of interaction between the lexemes of the lexicon and the rules of grammar. However, the term *vokabula* was not included in the “Conceptual apparatus” of (Apresjan, 2009). It can be found later in (Apresjan, 2014, pp. 7–8):

“*Vokabula* and *leksema*

The basic unit of description in the AS [= *Aktivnyj slovar' russkogo jazyka* / *Active dictionary of Russian* – T.R.], as well as in any other explanatory dictionary, is a single word, or a *vokabula*. [...]⁴.

The second very important unit of the dictionary is the word in one of its inherent senses, or the *leksema*. More precisely, a *leksema* in the AS is understood as a word taken in one of its senses and with the totality of the linguistically essential properties inherent to the word in this sense. Linguistically essential properties are properties on which some rules of a given language are based. [...]. They include the stylistic properties of a *leksema*, its phonetic properties, the specific prosody of the *leksema*, its grammatical forms, if they differ from the grammatical paradigm of the whole word, the possible meanings of grammatical forms, the *dictionary meaning* (*tolkovanie*) of the *leksema*, its regular modifications in certain contextual conditions, some information about the realities behind the given *leksema*, the government properties of the *leksema*, [...], non valency-controlled syntactic constructions characteristic of the *leksema*, its lexical-semantic combinability and,

---

<sup>4</sup> Speaking about the *vokabula* as the basic (!) unit of description in the *Active Dictionary of the Russian language*, Apresjan contradicts his own words („It is the *leksema*, not the word, that is the real lexical unit of the language”, see above) and Melčuk's approach (see below). For more details see also Footnote 12.

finally, its lexical world, i.e., synonyms, analogues, conversions, antonyms and derivatives.”<sup>5</sup>

As can be seen, Apresjan provides a lot of information about the concept of *leksema*. In contrast, the term *vokabula* remains rather underdefined, as does the related term *word* (Rus. *slovo*).<sup>6</sup>

## 1.2. Spencer & Wiemer (2020): *lexeme* and “*vocable*” in *ESLL*

For the sake of comparison, and to demonstrate the international recognition of the approach of Apresjan and Melčuk (for the latter, see Section 1.3), the following two paragraphs are taken from Spencer & Wiemer's (2020) recent encyclopedic article on “Lexeme” in the *ESLL* (Greenberg & Grenoble, 2020–):

“Like Zaliznjak (1967), Apresjan and his collaborators conceive of lexemes as abstract units represented by word forms or phrases (i.e., analytical/periphrastic expressions), called “lexes”, organized in paradigms (Melčuk, 2006, p. 21). Lexemes are opposed to “vocables” (Rus. *vokabuly*). The latter in practice correspond to dictionary entries (as defined, e.g., in Katamba 1993, pp. 17f.): a polysemic vocable is a *signifiant* that comprises as many lexemes as the number of meanings

<sup>5</sup> The original Russian text reads as follows: «Вокабула и лексема. Основной единицей описания в АС, как и во всяком другом толковом словаре, является отдельное слово, или вокабула. [...] Второй важнейшей единицей словаря считается слово в одном из присущих ему значений, или лексема. Более точно, под лексемой в АС понимается слово, взятое в одном из имеющих у него значений, но во всей совокупности присущих ему в этом значении лингвистически существенных свойств. Лингвистически существенными считаются свойства, на которые опираются какие-то правила данного языка. [...] К ним относятся стилистические свойства лексемы, ее фонетические свойства, специфическая просодия лексемы, ее грамматические формы, если они отличаются от грамматической парадигмы всего слова, возможные значения грамматических форм, словарное значение (толкование) лексемы, его закономерные модификации в определенных контекстуальных условиях, некоторые сведения о стоящих за данной лексемой реалиях, управляющие свойства лексемы, [...], характерные для лексемы невалентные синтаксические конструкции, ее лексико-семантическая сочетаемость и, наконец, ее лексический мир, а именно синонимы, аналоги, конверсивы, антонимы и дериваты».

<sup>6</sup> In Meaning ⇔ Text Theory. “word” is not a term, or at least should not be used as such, due to its ambiguity: 1. polysemous word (= vocable in MTT terminology) vs. 2. lexeme (Fr. *acception d'un vocable*; Eng. *wordsense* in MTT terminology) vs. 3. wordform (element of a lexeme, one of its *lexes* in MTT terminology). The term “vocable” has been put forward precisely in order to avoid this ambiguity. For a recent discussion of the term *word* in linguistic typology, see Haspelmath, 2023.

[read: senses!<sup>7</sup> – T.R.] defined by the metalinguistic paraphrases; the paraphrases have to reveal how the meanings are connected (Melčuk, 1995c, p. 500 n3; 1997, p. 346; see Polysemy). From this it follows that homonymous units belong to different vocables. This is summarized in the following definition:

A *lexeme* is a minimal unit of lexicographic description; it corresponds to a dictionary entry. A lexeme is a word in a specific sense, with all its syntactic and morphological peculiarities (characteristic precisely of this sense). A vocable is a family of lexemes such that: (i) their signifiers are identical; (ii) their signifiers [read: signifieds! – T.R.] contain a non-trivial common component. Two lexemes belonging to the same vocable are in relation of *polysemy*; two lexemes having identical signifiers but belonging to two different vocables are in relation of *homonymy*. (Melčuk, 1995a, p. 132 n1; see also Melčuk, 1974, p. 141 n1; 1995b, pp. 415f.; 1997, pp. 102f., 329; 2006, p. 21; Melčuk & Zholkovsky, 1995, p. 51 n2).<sup>8</sup>

Obviously, Spencer & Wiemer (2020) use the English term *lexeme* in a natural way as the equivalent of the Russian term *leksema*, while putting “vocable” in quotation marks before using it – in the second paragraph – as the plain English equivalent of the Russian term *vokabula*. The authors also link the terms *lexeme* and *vocable* to the notion of *dictionary entry*: a *vocable* is said to correspond “in practice” to a dictionary entry, while a *lexeme* is said to be “the minimal unit of lexicographic description”.<sup>8</sup>

<sup>7</sup> In Meaning ⇔ Text Theory, as generally in linguistics, different meanings of a polysemic word are called (*word*) *senses*. On behalf of “lexes” see Footnote 13 below.

<sup>8</sup> It is worth noting that Spencer & Wiemer use the three terms (*signifié – meaning – sense*) synonymously, and that they fail to mention a third component of the linguistic sign, the so-called *syntactics* (of the lexeme). Further on, Spencer & Wiemer (2020) discuss details that are not directly relevant to our exploration of this paper’s main topic. However, I believe they are interesting in their own right and worthy of a mention as a side note:

“Since a lexeme is a word (vocable) with a specific meaning, and paradigms are built on lexemes (not on vocables; [...]), we are left with the question how lexical meanings are identified and formulated. This is where most theories, and thus approaches to the lexeme, differ. According to the earlier tradition (cf. Reformatskij, 1967), the meaning of a lexeme is effectively just what is left when we eliminate the grammatical meanings from the meanings of wordforms. This presupposes that the grammatical meanings of a given language are known and that a distributional analysis of all wordforms in the language (including their morphemic structure and their syntactic distribution) is possible, allowing us to isolate sets of forms with an identical lexical meaning (e.g., Zaliznjak, 1967; Kempgen, 1981 on Russian; Laskowski, 1984, pp. 22–26; 1987; 1998, pp. 42–52 on Polish). What all these approaches have in common is that they start from inflection and paradigmatically organized wordforms (however, Laskowski, 1984; 1998 explicitly accounts for noninflected word-form classes) but do not appeal to lexical semantic representations.”

### 1.3. Mel'čuk: *lexie*, *article de dictionnaire*, *vocabole* in DEC and ESLL

Along with Russian, French is the best represented language in terms of lexicographic theory and practice of the Meaning  $\leftrightarrow$  Text Theory. See, for example, the four volumes of DEC, the *Dictionnaire explicatif et combinatoire du français contemporain*, i.e., the French version of the *Explanatory-Combinatorial Dictionary* (Mel'čuk et al. (1984, 1988, 1992, 1999)). The theoretical foundations were set out in Mel'čuk, Clas & Polguère (1995). As can be seen below, Mel'čuk, Clas & Polguère (1995, pp. 15–16) proposed the French term (*la*) *lexie* as a hyperonym of the Russian term *leksema*:<sup>9</sup>

“La *lexie* est l'unité de base de la lexicologie — en fait, son objet central et même, en schématisant un peu, son seul et unique objet. [...] Le concept de *lexie* est une formalisation et, simultanément, une généralisation de la notion de MOT. Il n'est pas dans notre intention d'entreprendre ici une analyse poussée de cette notion, même si l'on sait bien que le mot est une unité centrale de la langue: [...] on trouvera une étude détaillée de la question et des propositions de définitions dans Mel'čuk, 1993, Partie I, p. 97 *ssq.*”

Again, as before with Apresjan and Spencer & Wiemer, the understanding of the term *word* is taken to be intuitively clear and prior to the definition of the term *lexie*. Further on, *lexie* is revealed to be synonymous with *unité lexicale*, and hyperonymous to *lexème* and *phrasème*:

“Pour le moment, il nous suffit de dire qu'une *lexie* (n1: Nous utilisons le terme *lexie* tel que propose et défini par B. Pottier (1991)) ou *unité lexicale* est soit un mot pris dans une acception bien spécifique (= *lexème*), soit encore une locution, elle aussi prise dans une acception bien spécifique (= *phrasème*).”<sup>10</sup>

Later in their book, in Chapter IV on the macrostructure of the dictionary, Mel'čuk, Clas & Polguère (1995, pp. 155–156), explain the connection between the terms *lexie* and *article de dictionnaire*.

“L'unité de base du DEC, comme nous l'avons déjà dit, est la *lexie*; chaque *lexie* est décrite dans un article de dictionnaire spécifique. Une *lexie* ne correspond qu'à

<sup>9</sup> As we will see below, Fr. *lexie* (= Fr. *unité lexicale*) is a hyperonym of Fr. *lexème* (Rus. *leksema*) and of Fr. *phrasème* (Rus. *frazema*).

<sup>10</sup> A later term that was proposed instead of the too general *phrasème* is *locution = idiom*.

un seul article de dictionnaire, et *vice versa*, un article de dictionnaire correspond toujours à une seule lexie.”

In a further step, the French equivalent of the Russian term *vokabula* is established: the authors refer to it as (*le*) *vocable*, and to its corresponding lexicographic description as *le superarticle de dictionnaire*:

“On constate, cependant, que, dans de nombreux cas, des lexies ont un même signifiant et, en plus, manifestent entre elles des liens sémantiques assez évidents. Par exemple, on a en français au moins trois lexies ayant le signifiant /lapɛ̃/ qui sont sémantiquement liées: LAPIN (animal) (*Le lapin mange des carottes*), LAPIN (viande de cet animal) (*une excellente terrine de lapin*) et LAPIN (fourrure de cet animal) (*un manteau en lapin*). C’est le phénomène bien connu de la *polysémie* [...] Le DEC, [...] aura au moins trois articles différents pour les trois lexies LAPIN illustrées ci-dessus — trois articles différents réunis cependant sous le même vocable LAPIN. Dans le DEC, donc, un vocable n’est pas un article de dictionnaire mais un ensemble d’articles apparentés par leur sémantisme et leur forme — un *superarticle*.”

More recently, in their ESLL, article on the *Explanatory-Combinatorial Dictionary* Melčuk & Reuther (2020) used the term “vocable” in quotation marks. Building on the well-established term *lexical unit* (abbreviated LU), they introduced the term (*lexicographic*) *superentry* to refer to the collection of dictionary entries describing the *vocable*:

“The target of a lexical entry in the ECD is, as stated above, a “monosemous LU” – a lexeme or an idiom. Each LU is described by one dictionary entry, and each entry describes one LU. But what about “polysemous words,” which are so numerous in any language? This brings us to the notion of “vocable,” i.e., a lexicographic superentry: if a lexeme is analogous to a “word sense,” then a vocable corresponds to a “full word.”

A “vocable” is the set of all LUs such that

- (i) all of them have an identical signifier, and
- (ii) any one of them is semantically linked to all the others.”

The background for the quotation marks will be the subject of discussion in Section 2 below.

#### 1.4. Intermediate summary

Let us summarize what has been discovered so far.

- The term *leksema* as defined by Apresjan and Melčuk is a key concept in lexicology<sup>11</sup> and is directly linked to the corresponding concept of dictionary entry (Rus. *slovarnaja stat'ja*) in lexicography.
- In Apresjan's approach, "the basic unit of description in the AS is the vokabula.", whereas Melčuk holds that "l'unité de base du DEC est la lexie."<sup>12</sup>
- The Russian term *leksema* has equivalents in other languages, such as Eng. *lexeme*, Fr. *lexème*, while the Russian term *vokabula* (cf. Latin *vocabulum*) has evolved as a technical term in Russian linguistics and has been assigned terminological equivalents in French (*le*) *vocable* and in English (*the*) *vocable*.
- The French term (*la*) *lexie* has no generally accepted equivalent in other languages. Spencer & Wiemer (2020, see above) put "lexes" in quotation marks.<sup>13</sup>
- The French term *lexie* is said to be synonymous with Fr. *unité lexicale*; its equivalent in English is *lexical unit* (Rus. *leksičeskaja edinica*).
- The Russian term *slovarnaja stat'ja* has the commonly known equivalents *article de dictionnaire* in French and *dictionary entry* in English.
- The innovative French term (*le*) *superarticle (de dictionnaire)* has its equivalent (*lexicographic*) *superentry* in English, whereas there is no such established equivalent in Russian.<sup>14</sup>

<sup>11</sup> However, note that the older Russian term *leksiko-semantičeskij variant slova*, which was abandoned by the Moscow Semantic School in favor of *leksema*, can still be found in the works of influential Russian lexicographers.

<sup>12</sup> The difference in wording between Apresjan and Melčuk must not be overestimated. It stems from a difference in the perspective of linguistic modelling, which is evident in almost all of their work. Apresjan often "looks in the direction" of linguistic analysis from Text to Meaning. In the case of the dictionary, he starts with a word and all its possible usages, moving towards defining separate meanings, while keeping them under the "umbrella" of one *vokabula* as long as possible, i.e., for as long as polysemy can be argued for (homonymy being the alternative case for establishing two *vokabuly*). Melčuk's modelling always (!) proceeds in the direction of synthesis from Meaning to Text. This means moving from a semantic representation (of what the speaker wants to say) to identifying lexical units and grammatical relations that express this Meaning as a Text. In this approach, the dictionary contains all the necessary information about *lexical units*, which, under the condition of polysemy, come as a set called a *vocable*. However, both Apresjan and Melčuk share the concept of two major inherently connected components of linguistic description: the lexicon and the grammar.

<sup>13</sup> In Meaning ⇔ Text theory the term *lex* is defined as an 'element of a lexeme', and not as the English equivalent of *lexie*, which is called a *lexical unit* in English.

<sup>14</sup> One could think of *slovarnaja sverxstat'ja* or *slovarnaja superstat'ja*.

## Section 2: The multilingual perspective

In this section, we will examine the favorable and unfavorable conditions for the existence and/or emergence of equivalents of *leksema* and *vokabula* in four major European languages (English, French, German, Spanish).

### 2.1 Data from corpora and general dictionaries on Russian *vokabula* / *leksema* and German *Vokabel* / *Lexem*

All four words are attested in corpora of Russian and German, respectively, as well as in general dictionaries of the two languages.

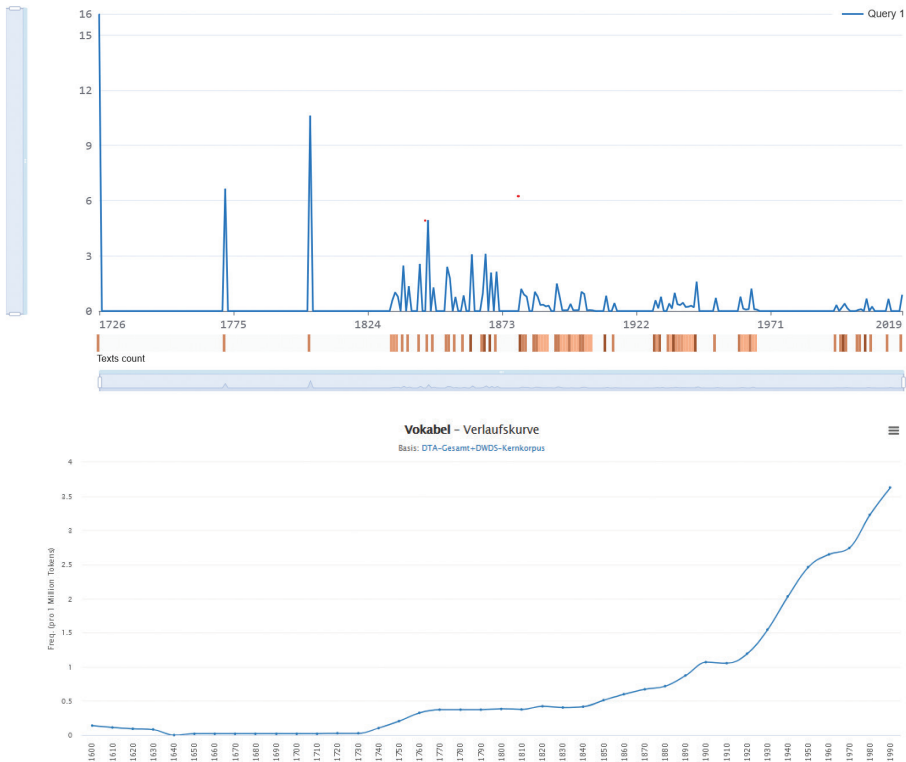
In Russian, the item *vokabula* is attested in the *National Corpus of the Russian Language* (ruscorpora.ru) and in general dictionaries, where it is found with one archaic sense (see below). As will become clear, this sense is different to that introduced by Apresjan and Mel'čuk for the lexicographic term.

In German, the item (*die*) *Vokabel* is equally attested in corpora — we will refer to Cosmas II corpus (cosmas2.ids-mannheim.de) — and is found in general dictionaries in two senses. In one of these senses, Ger. *Vokabel* is widely used in the teaching and learning of foreign languages, referring to a word that the student should learn, always in pair with an equivalent in the student's mother tongue (and usually written down for memorization and repetition in a special notebook called *Vokabelheft*). The question is whether there is enough “favorable space” for another language-related sense of *Vokabel*? Referring to corpus data and dictionary data, we will argue that there is not.

The following quantitative data were taken from the subcorpora specified as “Corpus name: main” in the *Russian National Corpus*, and *W Archiv der Geschriebenen Sprache; Korpus W-öffentlich - alle öffentlichen Korpora des Archivs W (mit Neuakquisitionen)*; *Archiv-Release: Deutsches Referenzkorpus DeReKo-2025-I* in the *Cosmas II corpus*:

Russian <i>vokabula</i>	Russian <i>leksema</i>	German <i>Vokabel</i>	German <i>Lexem</i>
Request 28.02.2025	Request 01.03.2025	Request 01.03.2025	Request 01.03.2025
Documents found: 61 Examples found 81	Documents found: 84 Examples found: 678	Examples found: 26.434	Examples found: 356

While *leksema* (678 examples found) and *Lexem* (356 examples found) appear to be closely related in Russian and German, the data for *vokabula* (81 examples found) and *Vokabel* (26434 examples found) are strikingly different in the two languages. The timelines of the Russian and German items provide more details: *vokabula* was popular in the 19th century, specifically from the 1820s to the 1870s, and is rarely used nowadays, whereas *Vokabel* has steadily increased in popularity from the 18th century to the present day.



Further on, in the dictionaries, the Russian word *vokabula* is monosemic and marked as archaic. The German word *Vokabel* is polysemic with one sense as provided above that is roughly equivalent to the archaic sense of the Russian word *vokabula*. Detailed lexicographic definitions<sup>15</sup> and examples from texts are given below:

<sup>15</sup> The original definition of Russian *vokabula*, taken from the four-volume *Small Academic Dictionary of Russian*, reads as follows: “Устар. Иностранное слово с переводом на родной язык как предмет заучивания при обучении какому-л. языку.” (<http://feb-web.ru/feb/mas/mas-abc/default.asp>). The German definition of *Vokabel* is taken from the online resource *Der deutsche Wortschatz von 1600 bis heute* (<https://www.dwds.de/wb/Vokabel>).

	<i>vokabula</i>	<i>Vokabel</i>
definition	<i>Archaic.</i> A foreign word with a translation into a person's mother tongue as an object of memorisation in foreign language teaching and learning	1. Einzelnes Wort einer fremden Sprache 2. Bezeichnung, Ausdruck
examples	<i>Učen'e vokabul, grammatiki i razgovorov šlo samo po sebe v klasse; no doma ja ničego ne učil naizust'</i> 'The study of <b>vocabulary</b> , grammar, and conversations went on by itself in class; but at home I learned nothing by heart' [S. Aksakov, 1856]; <i>Čto èto ty francuzskie vokabuly učiš'? — kivnul Ivan na tetradku, ležavšuju na stole</i> 'So you're learning <b>French word lists</b> ? — Ivan nodded toward the notebook lying on the table' [F. Dostoevskij, 1880].	1. <i>Bei solch' tollen Temperaturen <b>Vokabeln</b> pauken?</i> [Nürnberger Nachrichten]; <i>Eine neue <b>Vokabel</b> aus der Wirtschaftswelt ist in Rußland in aller Munde, obwohl noch die wenigsten etwas mit ihr anzufangen wissen: der Voucher, auch Privatisierungsscheck genannt.</i> [Süddeutsche Zeitung]; <i>Dabei wird neben Grammatik und <b>Vokabeln</b> auch der kulturelle Hintergrund unterrichtet.</i> [News]. 2. <i>"Katastrophal" – das war die meistgebrauchte <b>Vokabel</b> unter den Aktienanalysten, als sie die Gewinnprognosen des Siemens-Konzerns für das gerade erst begonnene Geschäftsjahr 1996/97 vernahmen.</i> [Nürnberger Nachrichten]; <i>Nicht eben das, was man sich unter der <b>Vokabel</b> "Sofortvollzug" vorstellt, den die Stadtentwässerung fordert.</i> [Süddeutsche Zeitung]; <i>Das Wort "löschen" ist die meistgebrauchte <b>Vokabel</b> der theatralischen gesellschaftlichen Bestandsaufnahme "Letzter Aufruf".</i> [News].

As a result, general dictionaries report the archaic Russian monosemic item *vokabula* and its polysemic German phonic equivalent *Vokabel*. The latter has two senses: *Vokabel 1*, which is related to language teaching and learning and is similar in meaning to *vokabula*, and *Vokabel 2*, which means 'denomination, expression'.

Adding the Russian linguistic term, as defined above by Apresjan and Mel'čuk (A&M), as a second lexeme *vokabula 2* would result in a polysemic Russian word *vokabula* with both archaic and modern senses. The "semantic bridge" between these two senses is 'word of a language' (a proof of their polysemic nature). The rest of the two senses, however, is to a certain degree antonymous, cf. 'one word in one language with its translation into another language'

for *vokabula 1* versus 'set of formally identical and semantically closely related lexemes of one language' for *vokabula 2*.

For the German word *Vokabel*, A&M's meaning is not attested, and I believe that there is an unfavorable condition to establish, along with *Vokabel 1* and *Vokabel 2* as defined above, a third lexeme *Vokabel 3* with A&M's terminological meaning. *Vokabel 1* is so well known and frequently used that another language-related, and to a certain degree antonymous (see above for *vokabula 2*) lexical unit *Vokabel 3* would confuse many general language users, as well as specialists. However, the remaining problem is how to express the meaning of A&M's term *vokabula 2* in German.

After turning to French and Spanish, as opposed to English, an alternative for both German and English will be given in Section 4 below.

## 2.2 French *vocable*, Spanish *vocablo*, English *vocable*

In this subsection we will demonstrate how French and Spanish differ from English in terms of the acceptability of the phonetic equivalents of *vokabula 2*.

### 2.2.1 French *vocable*

Checking authoritative online dictionaries of French, the *Larousse*, the *Robert*, and the *Oxford French Dictionary* reveals the following:

*Larousse en ligne* <https://www.larousse.fr/dictionnaires/francais/vocable/82341>)

1. Mot, terme d'une langue considéré surtout sous le rapport de sa signification, de son individualité lexicale (*petit, petite, petits* sont trois formes d'un même *vocable*).
2. Nom du saint ou du mystère sous le patronage duquel une église est placée.

*Robert en ligne* <https://dictionnaire.lerobert.com/definition/vocable>

Mot d'une langue, considéré dans sa signification, sa valeur expressive.

*Oxford French Dictionary online*

1. (mot) term;
2. *Religion* chapelle sous le *vocable* de chapel dedicated to.

As can be seen, all definitions refer to the hyperonym *mot* (word), and the meaning of *vocable 1* in *Larousse* is the closest to A&M's lexicographic term *leksema*. This is a favorable precondition for the integration of A&M's term as a sense of *vocable* into the French language.

### 2.2.2 Spanish *vocablo*

For Spanish, the reference sources are three dictionaries, the *Spanish Wiktionary online*, the *Diccionario del estudiante Real Academia Española online*, and the *Oxford Spanish Dictionary online*:

*Spanish Wiktionary online* <https://es.wiktionary.org/wiki/vocablo>

Unidad mínima del lenguaje escrito o hablado que puede darse separadamente

*Diccionario del estudiante Real Academia Española online* <https://www.rae.es/diccionario-estudiante/vocablo>

Palabra (sonido o conjunto de sonidos dotados de significado que constituyen una unidad indivisible del discurso, y su representación gráfica). *El vocablo “fútbol” proviene del inglés “football”. ¿Cuál es el significado de este vocablo?*

*Oxford Spanish Dictionary online*

(formal) word *vocablos extranjeros*

As can be seen, the Spanish item *vocablo* is a synonym of *palabra* (word). In the *Diccionario del estudiante Real Academia Española online*, there is a quasi-linguistic definition of what a *word (palabra)* is: “Word (sound or set of sounds with a meaning that constitute an indivisible unit of speech, and its graphic representation)”. This fact is a favorable condition for the integration of A&M’s term as a sense of *vocablo* into the Spanish language.<sup>16</sup>

### 2.2.3 English *vocable*

The *Oxford dictionary* (for British English) and the *Merriam Webster Dictionary* (for American English) both refer to the hyperonym *word*:

*Oxford dictionary of English online*

a word, especially with reference to form rather than meaning.

---

<sup>16</sup> Indeed, Alonso Ramos (2004), in her book related to the framework of the *Explanatory-Combinatorial Dictionary*, defines *vocablo* as a lexicographic term in line with A&M’s definition (“*vocablo* or polysemic word”): “Así, un *vocablo* o *palabra* polisémica como *dar* agrupará *lexemas* que tienen un comportamiento libre, es decir, que son escogidos según su sentido, y otros *lexemas* que son restringidos, es decir, que son controlados por otros ya escogidos previamente en el discurso. En otras palabras, un *vocablo* puede estar constituido por *lexemas* autónomos o libres y por *lexemas* colocativos.”

*Merriam-Webster online* <https://www.merriam-webster.com/dictionary/vocable#dictionary-entry-1>

term, specifically: a word composed of various sounds or letters without regard to its meaning. *The birds' phrasings are both melodic and mechanical, cyclical and spontaneous, like the wordless vocables of scat singers* [New York Times, 11 Feb. 2020].

As can be seen, both dictionaries offer only one meaning, which – although referring to *word* – is much narrower than the meaning of A&M's lexicographic term. In English, *vocable* is exclusively related to the form of a word, i.e., to its signifier, and not to the signified. This is an unfavorable condition for the integration of the word *vocable* in A&M's sense into the English language. Nevertheless, it is found as a term in quotation marks (Melčuk & Reuther, 2020, see above). An alternative name will be presented below.

To conclude this section, it became clear that the French *vocable* and the Spanish *vocablo*, both of which refer to *word/mot/palabra* as hyperonyms, can easily be used as lexicographical terms: *word/mot/palabra* refer to form and meaning, the traditional dual nature of lexical items. This creates favorable conditions for the integration of A&M's term into French and Spanish. On the contrary, the definition of *vocable* in English does not include the signified component of *word* (cf. 'a word composed of various sounds or letters without regard to its meaning'), resulting in unfavorable conditions for the integration of *vocable* in A&M's sense into the English language.

### Section 3.

#### **A short exploration of the terminology of traditional, computational and corpus linguistics: *lemma*, *keyword***

In order to solve the problem of how the meaning of A&M's term *vokabula* can be expressed in English and German, let us briefly consider two examples of terminology from computational and corpus linguistics.

Firstly, the term *keyword* has been used in lexicography for centuries, as has its German equivalent *Stichwort*. However, until now, neither *keyword* nor *Stichwort* have been used as equivalents of Russian *vokabula*.

Secondly, the recently updated version of the Russian National Corpus uses the term Rus. *lemma* (written in Cyrillic as *лeмма*) for the top search box in

the Russian user interface in parallel with Eng. *lemma* for the top search box in the English user interface. This demonstrates the complete integration of the term *lemma* into both Russian and English. However, *lemma* has not yet been suggested as a potential equivalent to *vokabula* in English.

In the next section, we will suggest German and English terminological equivalents of the lexicographic term Rus. *vokabula* (Fr. *vocable*, Sp. *vocablo*) in A&M's sense which make use of Ger. *Stichwort* (alternatively, *Lemma*) and Eng. *keyword* (alternatively, *lemma*).

#### **Section 4: A traditional and a corpora-oriented proposal for equivalents of Apresjan & Mel'čuk's term *vokabula/vocable* in German and English**

Obviously, the term *Stichwort* can be used to refer either to one lexeme, or to the set of lexemes that fall "under the umbrella" of a *vokabula* in the sense of A&M's term, cf. the following example: "Zum <Unter dem> *Stichwort* HASE findet man im Wörterbuch drei Bedeutungen." (Rus. "Vokabula HASE imeet v slovare tri znachenija." – for English, see below). Consequently, when it comes to the term *dictionary entry*, the term *Stichworteintrag* can be used referring to the "name of the umbrella" where all information about one or more lexemes that fulfil the definition of A&M's term *vokabula* is stored. Cf. "Der *Stichworteintrag* HASE enthält drei Unterartikel zu drei verschiedenen Bedeutungen des Wortes Hase." In other words, a *Stichworteintrag* is a dictionary entry for either a monosemic or for a polysemic word.

What remains to propose is a more favorable name for A&M's notion of *vokabula* in English. In my opinion, this could be either the traditional term *keyword* or the international term of Greek origin Eng. *lemma* (Rus. *lemma*, Fr. *lemme*, Ger. *Lemma*, Sp. *lema*, etc.).

If we agree to this solution, the quotation from Mel'čuk & Reuther (2020), cited in Section 1, Subsection 1.3, would read as follows:

"The target of a lexical entry in the ECD is, as stated above, a "monosemous LU" – a lexeme or an idiom. Each LU is described by one dictionary entry, and each entry describes one LU. But what about "polysemous words," which are so numerous in any language? This brings us to the notion of *keyword* (or *lemma*), i.e.,

a lexicographic superentry: if a lexeme is analogous to a “word sense,” then a keyword (or lemma) corresponds to a “full word.”

A keyword (or lemma) is the set of all LUs such that

- (i) all of them have an identical signifier, and
- (ii) any one of them is semantically linked to all the others.”

An alternative approach is to do without *keyword* and *lemma* and stick to (*lexicographic superentry*), cf.: “... if a lexeme is analogous to a “word sense,” then a superentry corresponds to a “full word.”

A superentry is the set of all LUs such that

- (i) all of them have an identical signifier, and
- (ii) any one of them is semantically linked to all the others.”

Consequently, the translations of the German sentences above are the following: “*In the dictionary, the keyword <lemma, superentry> HASE has three meanings.*”; “*The superentry HASE contains three subentries for three different meanings of the word Hase.*”

That concludes my contribution as a linguist and a native speaker of German. It is for linguists and native speakers of English to decide on the solution for English.

Our findings presented in Sections 2 to 4 are as follows:

1. Firmly anchored in Apresjan's and Mel'čuk's lexicographic theory and practice, the Russian term *leksema* and its lexical parallels (Eng. *lexeme*, Fr. *lexème*, Ger. *Lexem*, Sp. *lexema*) have made their way into international lexicographic terminology.
2. The Russian term *vokabula* can easily be integrated into the French and Spanish linguistic terminology but meets resistance in German and English. Luckily enough, Ger. *Stichwort* and Eng. *Lemma*, in our opinion, fulfill favorable conditions for their integration as equivalents of A&M's Russian term *vokabula* into German and English. In combination with the terms Ger. (*Wörterbuch*)*eintrag* and Eng. (*dictionary*) *entry*, we consider Ger. *Stichworteintrag* and Eng. *superentry* as favorable terms for the complete information for all lexemes that fall “under the umbrella” of a Rus. *vokabula*, Ger. *Stichwort*, Eng. *lemma*.

## Conclusion

Given the data on Russian, English, French, German and Spanish, the terms Rus. *leksema* – Eng. *lexeme* – Fr. *lexème* – Ger. *Lexem* – Sp. *lexema* are full terminological lexical parallels. On the contrary, the words Rus. *vokabula* – Eng. *vocabulary* – Fr. *vocabulaire* – Ger. *Vokabel* – Sp. *vocablo* have a historical background, dating back centuries before their terminological sense was first introduced by Apresjan and Melčuk for Russian. The lexical parallels of Rus. *vokabula* – Fr. *vocabulary* – Sp. *vocablo* proliferated easily into French and Spanish terminology, leaving the English equivalent “*vocabulary*” halfway in quotation marks, and the German *Vokabel* undocumented.

The situation in English and German is the following: In modern English, *vocabulary* is defined as ‘a word, especially with reference to form rather than meaning’. In modern German, *Vokabel* is defined as a one-to-one correspondence between two words, one in a foreign language, and one in a learner’s mother tongue. Both definitions contradict the idea of *vokabula* as a set of lexemes in one and only one language.

In order to integrate the concept of *vokabula* into German and English terminology, the following is proposed: The terms Ger. **Stichwort** (alternatively, **Lemma**), and Eng. **keyword** (alternatively, **lemma**) can refer either to one lexeme, or to the set of lexemes that fall “under the umbrella” of a *vokabula* in the sense of Apresjan & Melčuk’s terminology; consequently, Ger. *Stichwort* or *Lemma* and Eng. *keyword* or *lemma*, should be used as equivalents of the terms Rus. *vokabula* – Fr. *vocabulary* – Sp. *vocablo*. Further on, Ger. **Stichworteintrag** (alternatively, **Lemmaeintrag**) and Eng. **superentry** (as proposed by Melčuk) can be used, referring to the “name of the umbrella” where all information about one or more lexemes that fulfil the definition of A&M’s term *vokabula* is stored.<sup>17</sup>

## References

- Alonso Ramos, M. (2004). *Las construcciones con verbo de apoyo*. Visor Libros, Madrid.  
 Apresjan, Ju. (1986). Integral’noe opisanie jazyka i tolkovyj slovar’. *Voprosy jazykoznanija*, 2, 57–70.

---

<sup>17</sup> Many thanks to the reviewers for their comments on the first version of this paper, and to Igor Melčuk for his feedback and clarification of his position on several drafts of this paper. The remaining views on terminology are my own.

- Apresjan, Ju. (2009). Ponjatijnyj apparat sistemnoj leksikografii, In Ju. Apresjan, *Issledovanija po semantike i leksikografii. Tom 1. Paradigmatika* (p. 486–552). Jazyki slavjanskix kul'tur, Moskva.
- Apresjan, Ju. (2014). Ob aktivnom slovare russkogo jazyka. In Ju. Apresjan (Ed.), *Aktivnyj slovar' russkogo jazyka. Tom 1. A-B* (pp. 5–36). Jazyki slavjanskoj kul'tury, Moskva.
- Dubichynskij, V., & Reuther, T. (2020). *Leksygrafuvannja leksyčnyx paralelej/Lexikografie lexikalischer Parallelen. Teoretyčni poležennja ta ukrajins'ko-nimec'kyj slovnyk. Monografija / Theoretische Grundlagen und ukrainisch-deutsches Wörterbuch. Monografie*. Peter Lang, Berlin.
- Greenberg, M. L., Grenoble, L. A., & others (Eds.). (2020–). Encyclopedia of Slavic Languages and Linguistics Online. Leiden & Boston: Brill. <https://doi.org/10.1163/2589-6229>
- Haspelmath, M. (2023). Defining the word. *Word, Vol. 69, No. 3*, 283–297, <https://doi.org/10.1080/00437956.2023.2237272>
- Katamba, F. (1993). *Morphology*. Macmillan, London.
- Kempgen, S. (1981). "Wortarten" als klassifikatorisches Problem der deskriptiven Grammatik. *Historische und systematische Untersuchung am Beispiel des Russischen*. Otto Sagner, München.
- Laskowski, R. (1984). Podstawowe pojęcia morfologii. In R. Grzegorzczkova et al., *Gramatyka współczesnego języka polskiego, vol. I: Morfologia* (pp. 9–57). Wydawnictwo Naukowe PWN, Warszawa.
- Laskowski, R. (1987). On the concept of the lexeme. *Scando-Slavica*, 33, 169–178.
- Laskowski, R. (1998). Zagadnienia ogólne morfologii. In R. Grzegorzczkova et al., *Gramatyka współczesnego języka polskiego, vol. I: Morfologia*. (pp. 27–86). Wydawnictwo Naukowe PWN, Warszawa.
- Melčuk, I. (1974). *Opyt teorii lingvističeskix modelej "Smysl ⇔ Tekst"*. Nauka, Moskva.
- Melčuk, I. (1993). *Cours de morphologie générale (théorique et descriptive)*. Vol. 1. *Introduction et Première partie: Le mot*. Les Presses de l'Université de Montréal – CNRS, Montréal.
- Melčuk, I. (1995a). Countability vs. non-countability of nouns in Russian and their lexicographic description. In I. Melčuk, *Russkij jazyk v modeli "Smysl ⇔ Tekst"* (pp. 125–133). Jazyki russkoj kul'tury – Wiener Slawistischer Almanach.
- Melčuk, I. (1995b). La combinatoire des numéraux du type DVOE en russe contemporain. In I. Melčuk, *Russkij jazyk v modeli "Smysl ⇔ Tekst"* (pp. 383–422). Jazyki russkoj kul'tury – Wiener Slawistischer Almanach, Moskva – Wien.
- Melčuk, I. (1995c). Slovoobrazovanie v lingvističeskix modeljax tipa Smysl-Tekst (predvaritel'nye zamečanija). In I. Melčuk. *Russkij jazyk v modeli "Smysl ⇔ Tekst"* (p. 475–504). Jazyki russkoj kul'tury – Wiener Slawistischer Almanach, Moskva – Wien.

- Melčuk, I. (1997). *Kurs obščej morfologii. Tom I. Čast' 1: Vvedenie. Slovo*. Jazyki ruskokul'tury, Wiener Slawistischer Almanach, Moskva – Wien.
- Melčuk, I. (2006). *Aspects of the theory of morphology*. Walter de Gruyter, Berlin–New York.
- Melčuk, I., et al. (1984, 1988, 1992, 1999). *Dictionnaire explicatif et combinatoire du français contemporain: Recherches lexico-sémantiques (I–IV)*. Les Presses de l'Université de Montréal, Montréal.
- Melčuk, I., Clas, A., & Polguère, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. Éditions Duculot, Louvain-la-Neuve.
- Melčuk, I., & Reuther, T. (2020). *Explanatory Combinatorial Dictionary*. In M. L. Greenberg, L. A. Grenoble et al. (Eds.), *Encyclopedia of Slavic Languages and Linguistics Online*. Brill <https://doi.org/10.1163/2589-6229>
- Melčuk, I., & Zholkovsky, A. (1995). *An explanatory combinatorial dictionary of modern Russian*. In I. Melčuk, *Russkij jazyk v modeli "Smysl ↔ Tekst"* (pp. 17–54). Jazyki ruskokul'tury – Wiener Slawistischer Almanach, Moskva–Wien.
- Plungjan, V. (2000). *Obščaja morfologija: Vvedenie v problematiku*. URSS, Moskva.
- Pottier, B. (1991). La lexie: une mise au point nécessaire. In A. Zampolli (Ed.), *Computational lexicology and lexicography. Special issue dedicated to Bernard Quemada*. Giardini, Pisa.
- Reformatskij, A. (1967). *Vvedenie z jazykovedenie*. 4-e izdanie. Prosveščenie, Moskva.
- Spencer, A., & Wiemer, B. (2020) *Lexeme*. In M. L. Greenberg, L. A. Grenoble et al. (Eds.), *Encyclopedia of Slavic Languages and Linguistics Online*. Brill, Leiden, Boston.
- Zaliznjak, A. (1967). *Russkoe imennoe slovoizmenenie*. Nauka, Moskva.