# Exclusion Labels in Slavic Monolingual Dictionaries

## Lexicographic construal of non-standardness

Danko Šipka*

## Abstract

The present paper analyzed the practices of excluding vocabulary items from the standard language in four major Slavic monolingual dictionaries (Polish, Russian, Serbian, and Slovenian). The analysis has revealed that there exist a range of primary (overt) and secondary (implicit) exclusion labels and that their effect is negotiated in the sociocognitive context of the dictionary, in a communication between the dictionary and its users.

*Key words:* standard language, usage labels, linguistic norms, cognitive sociolinguistics, metalexicography

## 1 Theoretical Background and Conceptual Map

The term exclusion labels, as used here, is a subset of usage labels. We will hence commence with defining usage labels. In its broader sense, *usage* refers to any kind of relationship of words or their features (their meanings or their forms) and their language production/reception context. This ranges from the required morphosyntactic form of that particular context to semantic links the word may have, and the effects the word creates in that context. In its narrower sense, *usage* excludes any phonetic, phonological, prosodic, morphosyntactic, and core semantic parameters. In that sense, this term refers

---

* Arizona State University, Department of Languages and Literatures; danko.sipka@asu.edu

to a higher, lower, or absolute valence of the word or its feature toward certain contexts (e.g., toward one period in time or region but not toward another) or a certain contextual effect (e.g., offensiveness, facetiousness, etc.). We will employ *usage* in this narrower sense and the concrete scope of the phenomena covered with it will be outlined in the second section of this paper based on the review of the practice in lexicography and metalexicography. *Label* is any succinct, customary and, as a rule, repeatedly deployed reference to the usage parameters. In a typical case, such customary and repeated references are either listed in the front matter, e.g., *obs.(olete)*, *Am.(erican)*, *off.(ensive)*, or established in linguistic and/or lexicographic practice of the language in question (e.g., grammatical labels showing the inflection of the word, such as *gave*, *given* under *give*). Not every reference to usage is a label – there are also glosses, which refer to usage, such as *used only in Shakespeare's works* for the word *honorificabilitudinitatibus*. *Usage labels*, given that the narrower meaning of usage is employed here, refer to succinct customary references to the valence of the words or their features toward certain contexts or contextual effects. This precise usage of the term can be found in Béjoint (2010, 12): "A usage label is a noun or adjective indicating the kind of context in which the word is normally used: *slang*, *literary*, *American*, *Medicine*, etc. Usage labels are usually abbreviated (*sl*, *lit*, *US*, *med*, etc.). They are of different kinds, social, geographical, stylistic, etc. corresponding to different varieties of language".

Under exclusion labels we understand a kind of usage labels which have the potential of expressing the attitude of exclusion of a labeled item from contemporary standard language or at least its core. The idea of standard and non-standardness will be discussed in the second section of this paper and the labels themselves will be addressed in its third section.

The present research is informed by cognitive sociolinguistics and communicative metalexicography. The following two statements are relevant in this regard.

> As natural as it is for Cognitive Linguistics to study the variation of meaning, is it just as natural to study the meaning of variation, i.e. the way in which language users make sense of linguistic variation, the way in which linguistic variation is meaningful to them. In a usage-based conception of language, we assume that language users have a cognitive representation of the communicative interactions in which they participate: that – rather than some genetic endowment – is their 'knowledge of the language'. But as their interactive horizon includes linguistic variation, they also have a representation of that diversity. They categorize social reality as reflected in language use and differences of language use, and such a categorization process is typically one of the phenomena that Cognitive Linguistics is interested in: 'meaning as categorization' is a kind of catchphrase in Cognitive Linguistics. The questions that arise here are of the following kind.

– How do language users perceive lectal differences, and how do they evaluate them attitudinally?
– What models do they use to categorize linguistic diversity?
– How does linguistic stereotyping work: how do language users categorize other groups of speakers?
– What is the role of subjective and objective linguistic distances: is there a correlation between objective linguistic distances, perceived distances, and language attitudes?
– Are there any cultural models of language diversity: what models of lectal variation, standardization, and language change do people work with?
– To what extent do attitudinal and perceptual factors have an influence on language change?
– How do language users acquire lectal competence, how is it stored mentally, and how does it work in language production?

(Geeraerts et al. 2010, 9–10)

The present study attempts to establish a triangular communicative model of lexicography and views the bilingual dictionary as a system of intercultural communication between the compiler and the user. (Yong and Peng 2007, 15)

If we apply the two aforementioned approaches to the study the dictionary labels as exclusion tools of standardization, we first need to address the idea of non-standardness (expounded in section two) followed by the role of dictionary labels in the context of declaring a lexical item, its form or meaning as substandard (discussed in section three).

## 2  The Idea of Non-standardness

The idea of non-standardness assumes the existence of a standard language variety. In defining what that standard language variety is, we will start with the social dynamic field of a standard variety (Ammon 1995, 80, summarized in English in Takahashi 2004, 177–178). This field, as defined by Ammon, includes five key elements: majority population (i.e., the members of the community of speakers), norm authorities, in charge of language corrections, codifiers, who create language codex (such as orthographic dictionaries), linguistic experts, who issue their expert opinions, and model speakers/and writers, whose linguistic performance serves as the basis for norm codification. Majority population creates primary rules, which are uncodified, while four other aforementioned factors of the process create secondary norms, which are codified.

Ammon (2004) elaborates on this process and goes on to note that "[...] the categorization (delimitation) of standard from non-standard forms is not always a simple

yes-no decision, though there are unquestionable standard forms and unquestionable non-standard forms, namely those for which all the abovementioned forces agree" (Ammon 2004, 278, reference is made to majority population, codifiers, linguistic experts, norm authorities, and model speakers/writers).

Monolingual dictionaries of standard Slavic languages typically involve all four factors identified by Ammon: they attract considerable interest in majority population, lexicographers are seen as codifiers, linguist experts are present and so are norm authorities, and finally dictionary examples typically involve model writers. Despite the fact that dictionary compilers may state that their dictionary is descriptive rather than normative, dictionary users (Ammon's majority population) perceive such dictionaries as a source of authority. As Milroy and Milroy (2012, 4) aptly state: "The attitudes of linguists (professional scholars of language) have little or no effect on the general public, who continue to look to dictionaries, grammars and handbooks as authorities on 'correct' usage. If, for example, lexicographers (dictionary-makers) attempt to remove all traces of value-judgment from their work and refuse to label particular usages (such as ain't) as 'colloquial' and others as slang, there is likely to be a public outcry." A similar thought can be found in Ammon (2004, 276–7) in his discussion of "social forces establishing and controlling standard varieties".

Needless to say, the ideas of what constitutes the standard language vary quite considerably, as demonstrated by Smakman (2012) who surveyed over one thousand non-linguists from England, Flanders, Japan, the Netherlands, New Zealand, Poland and the United States and concluded that although the ideas about the standard language have very little in common, one can recognize a socially distinctive (exclusive) notion of the standard language and a socially cohesive (integrative) one.

The role of monolingual dictionaries in standardization can be seen in (mostly political) attempts to form new Slavic standard languages. Thus, Dževad Jahić, a standardizer of Bosnian, states the following in a TV interview: "In order for a language to be normed, it has to have an international language code [...] as well as a grammar, a dictionary, and a manual of orthography, in that order" (https://www.youtube.com/watch?v=LXV9cRTMvjU, 29:00–29:17, accessed on June 29, 2015). Similarly, a recent manual of Montenegrin orthography features the following passage: "[...] the secretary of education and science [...] asked the Council [on Montenegrin language] to [...] form a task force charged with preparing a final version of a proposal for a manual of orthography, grammar, and dictionary of the Montenegrin language" (http://www.gov.me/files/1248442673.pdf, accessed on June 30, 2015)

Dittmar (2004) discusses the idea of nonstandard in German, English, and French speaking cultures. German tradition is definitely closest to the Slavic language traditions in particular in its establishment of colloquial language (Umgangssprache), which serves somewhat like a buffer zone between standard and substandard. Ammon (2004, 278) notes: "[...] *colloquial* (or German *umgangsprachlich*) [...] can be taken to mean

(1) 'colloquial standard', i.e., a stylistic level within the standard norm or (2) 'non-standard'"

The idea of standardness and hence non-standardness is rather problematic. Ammon (2004) offers an elaborate account of the level of ambiguity and category overlapping in this sphere. For our purposes here, we will understand standard language in its narrower sense, as those language varieties subject to standardizing efforts by codifiers, linguistic experts, and norm authorities. We are thus following Smakman's (2012) socially distinctive (exclusive) notion of the standard language, as this variety needs to be edited (using dictionaries as reference). This would include administrative, legal, scholarly language, news, etc. and exclude informal colloquial and literary language (which are typically not the target of normative interventions). This understanding of standard language is the best platform for the study of the role of exclusion labels as it is uncontroversial that a kind of lexical norm can be enforced in such varieties by excluding certain words from the norm. Needless to say, an operational definition of standard deployed here does not attempt to establish a new understanding of standard language – it is merely striving to simplify matters and make them analyzable.

In that view, non-standard vocabulary would include the items that are inappropriate in standard language varieties. Thus, for example excretion or sex-related obscenity would be non-standard given that it cannot be freely used in standard language varieties (unlike in colloquial and literary genres). Dictionaries would hence typically mark such word accordingly.

## 3 Exclusion Attitude Labels in the Context of non-standardness

What we named exclusion labels represents a subset of lexicographic usage labels. Svensén (2009, 316) provides a very comprehensive account of usage labels types as shown in Table 1.

One should note that Table 1 represents a handy English-language version of the categories found in Hausmann (1989, 651) with another version found in Hartmann and James (2001, 151). In this approach, subject-matter labels are called diatechnical, they are restricted to a specific technical language and contrasted to general language. One should say that this classification remains strongly Eurocentric. For example, it does not contain gender (words used exclusively by males or females, i.e., in texts marked as such), age (items used by younger or older people, etc.) and other similar categories, as these are not so frequent in European languages and rather rare in European dictionaries. (On a side note, one should state that even European dictionaries use labels such as "youth language".) The very idea of something being technical presumes a kind of society we know from European countries. In contrast, the idea of subject matter would

| Criterion | Type of marking | Unmarked centre | Marked periphery | Examples of labels |
|---|---|---|---|---|
| Time | diachronic | contemporary language | archaism – neologism | arch, dated, old use |
| Place | diatopic | standard language | regionalism, dialect word | AmE, Scott, dial. |
| Nationality | diaintegrative | native word | foreign word | Lat., Fr. |
| Medium | diamedial | neutral | spoken – written | colloq., spoken |
| Socio-cultural | diastratic | neutral | sociolects | pop., slang, vulgar |
| Formality | diaphasic | neutral | formal – informal | fml, infml |
| Text type | diatextual | neutral | poetic, literary | journalese, poet, lit. |
| Technicality | diatechnical | general language | technical language | Geogr., Mil., Biol., Mus. |
| Frequency | diafrequential | common | rare | rare, occas. |
| Attitude | diaevaluative | neutral | connoted | derog., iron., euphem. |
| Normativity | dianormative | correct | incorrect | non-standard |

Table 1: Types of usage labels (Svensén 2009, 316)

be much more appropriate for other cultural settings, given that it does not presume any kind of advanced technology.

A majority of established types of labels may serve as exclusion labels. The most direct labels of this kind are dianormative ones. Exclusion can also, rather straightforwardly, be achieved by diachronic, diatopic, and diastratic labels. All those labels exclude the word from contemporary standard language. A number of diaphrasic labels (e.g., informal) as well as diaevaluative labels certainly have an indirect exclusion potential. In other words, while their primary role is to mark the attitude, they carry a certain potential of being understood by the users as exclusion labels (see section 2 on the discrepancy between the compilers' intentions and users' understanding of the text). Other labels are just marking an area or a feature within standard language.

# 4  Non-standardness Labels in Slavic Lexicographic Traditions

Monolingual Slavic general dictionaries have a two-way communication with the society. On the one hand they are at the receiving end of normative authorities (i.e., they are expected to follow general delimitations of the standard and substandard sphere). On the other hand, lexicographic strategies, in particular the use of exclusion labels, are an agent of normative enforcement.

Slavic languages have a tradition of considering literary languages as standard. Most often, references to standard language already include the realm of literature. *Literaturnyj jazyk* in Russian, *język literacki* in Polish, *knjižni jezik* in Slovene, and *književni jezik* in Serbian all mean 'literary language'. Attitudes toward standard language are

perhaps best reflected in Ščerba (1957), where he states that the most basic function of literary language is to be generally accepted and generally understandable: "[...] ос-новную функцию, которую должен выполнять литературный язык и которая в сущности только и делает его литературным, т.е. **общепринятым**, а потому и **общепонятным**" (Ščerba 1957, 114). It is of particular importance that Ščerba offers linguistic arguments for a campaign initiated by Maxim Gorky to protect Russian literary language from being dirtied by dialectisms and vulgarisms: "беречь русский литературный язык от засорения его диалектизмами и вульгаризмами" (Ščerba 1957, 128). We can clearly see the idea of non-standardness, and the need to exclude certain lexical items from the standard.

Practically all Slavic lexicographic traditions follow this literary language agenda and have the need to define it by excluding certain items, which belong to the forms of language that are not generally accepted and/or generally understandable. We will now look into such lexical strata and their corresponding usage labels.

In practically all contemporary monolingual Slavic dictionaries (and four of them will be analyzed in the next section) one can find the following two ideal types of exclusion labels:

a) Primary exclusion labels, i.e., such labels as *slang*, where the dictionary compiler's intention was to clearly exclude the word, its meaning, or form from the standard language,

b) Secondary exclusion labels, e.g., *vulgar*, where the label was used to signal something else (in this case the attitude), but its effect in users can be that of exclusion.

In real-life application these ideal types blend into one another and there, in fact, exists a continuum of exclusiveness from strong and clear exclusion from the standard language to a weak and borderline exclusion effect.

To test the effect of exclusion, a list of labels was presented to the users of Serbian dictionaries (the labels were taken from Vujanić et al. (2011), a Serbian monolingual dictionary). Those labels that were considered potentially excluding from the standard language were included in this list. The participants were asked to determine if the word marked by such label would be appropriate for the use in standard language (e.g., in evening news on national TV). Most participants were recruited from the Facebook list "Naš jezik", which includes persons interested in the Serbian language, most of them users of Serbian monolingual dictionaries. The first nine labels were primary exclusion labels as they mark the word as belonging to a lexical subfield which is predominantly non-standard, or at least it does not belong to the core standard-language lexicon. The remaining six labels refer to the attitude associated with the word or its feature, so their primary role is not exclusion. The results were obtained as shown in Table 2.

As can be seen, all labels listed in this survey have a potential of excluding a lexeme or its feature from standard Serbian. While the effect of exclusion is indeed somewhat lower in attitude labels, there is a solid number of users (on average well over one third for

Choose if this label means that the word marked by it is non-standard

|  | yes | no |
|---|---|---|
| slang (шатровачки) | 78.26% (54) | 21.74% (15) |
| jargon (жаргонски) | 68.12% (47) | 31.88% (22) |
| regional (регионално) | 57.97% (40) | 42.03% (29) |
| vulgar (вулгарно) | 56.52% (39) | 43.48% (30) |
| old literary style (старокњижевно) | 56.52% (39) | 43.48% (30) |
| archaic (архаично) | 55.88% (38) | 44.12% (30) |
| folksy (народски) | 54.41% (37) | 45.59% (31) |
| school student speech (ђачки) | 51.47% (35) | 48.53% (33) |
| colloquial (разговорно) | 49.28% (34) | 50.72% (35) |
| familial/homely (фамилијарно) | 47.06% (32) | 52.94% (36) |
| pejorative (пејоративно) | 45.59% (31) | 54.41% (37) |
| derisive (подругљиво) | 34.78% (24) | 65.22% (45) |
| contemptuous (презриво) | 30.43% (21) | 69.57% (48) |
| ironic (иронично) | 27.54% (19) | 72.46% (50) |
| facetious (шаљиво) | 27.94% (19) | 72.06% (49) |

Table 2: User's judgements about acceptability of labels in the standard language

these labels) who perceive them as excluding. There are also such users whose definition of the standard language is so broad to include in it all spheres of the lexicon.

To eliminate the factor of different definitions of the standard language and give the decision about non-standardness a concrete touch and feel, another survey was conducted, asking if concrete words, labeled in Vujanić et al. (2011) with one of the aforementioned labels, would be acceptable in the speech of an evening news anchor. The following results were obtained:

| Word (label in Vujanić et al. 2011) | yes | no |
|---|---|---|
| говнарија (vulgar) | 0.00% (0) | 100.00% (40) |
| окинути на годину (school students speech) | 0.00% (0) | 100.00% (40) |
| жаца (ironic) | 0.00% (0) | 100.00% (40) |
| аратосиљати се (folksy) | 2.50% (1) | 97.50% (39) |
| бос по глави (facetious) | 2.50% (1) | 97.50% (39) |
| цвикераш (derisive) | 2.50% (1) | 97.50% (39) |
| гологуз (pejorative) | 2.50% (1) | 97.50% (39) |
| ачити се (colloquial) | 5.00% (2) | 95.00% (38) |
| брајкан (folksy and familial/homely) | 5.00% (2) | 95.00% (38) |
| гладибрк (facetious) | 5.00% (2) | 95.00% (38) |
| безец (jargon) | 7.50% (3) | 92.50% (37) |
| лацман (folksy) | 10.00% (4) | 90.00% (36) |
| наровашен (regional) | 12.50% (5) | 87.50% (35) |
| дажд (archaic) | 20.51% (8) | 79.49% (31) |

Table 3: User's judgements about acceptability of labeled words in the standard language

As can be seen very clearly, concrete words in a concrete setting are seen as generally unacceptable in standard language. It is also worth noting that there is no difference between primary and secondary exclusion labels. We will now analyze four dictionary case studies.

# 5  Case Studies

The following four case studies of contemporary Slavic monolingual dictionaries will be discussed: Russian (Ožegov and Švedova 1992), Polish (Żmigródzki et al. 2012), Serbian (Vujanić et al. 2011), and Slovenian (Bajec et al. 2000).

These four dictionaries share their general approach to standardness and non-standardness but they differ substantially in their lexicographic deployment of primary and secondary exclusion labels. These differences are seen even in the front matter, where there exists considerable variation in the space devoted to explaining the set labels (from the extremely elaborate Polish case to laconic Serbian). Then, each dictionary deploys a slightly different set of labels and labels its entries differently.

Each of the four dictionaries examined here will be analyzed first, which will be followed by a comparison of their particular approaches. For each of the four dictionaries we will first discuss the set of relevant labels and their description and then turn to their count.

## 5.1  Russian Monolingual Dictionary

Ožegov and Švedova (1992) provide a 612-word-long text about usage characteristics of the words and their marking in the front matter of their dictionary. The authors divide usage labels (they call them notes, "пометы") into three categories: a. stylistic, i.e., labels expressing style characteristics of the word ("пометы, указывающие на стилистическую характеристику слова"), b. specialized, i.e., technical use ("специальное"), and c. temporal labels, i.e., those that point to the historical context of the word's usage ("пометы, указывающие на историческую перспективу слова"). This approach is quite interesting in that most primary and secondary exclusion labels belong to the same category, i.e., to stylistic labels.

The following repertoire of relevant exclusion labels is deployed in this dictionary:

(a) Stylistic labels: non-literary colloquial ("просторечное"), regional ("областное"), contemptuous ("презрительное"), disapprobatory ("неодобрительное"), disparaging ("пренебрежительное"), facetious ("шутливое"), ironic ("ироническое"), vituperative ("бранное"), rude ("грубое")

(b) Temporal labels: old ("старое") and obsolete ("устаревшее")

It is interesting to note that there is also the category of colloquial ("разговорное"), which is defined as being a part of the standard language, just more informal ("оно не выходит из норм литературного словоупотребления, но сообщает речи непринуждённость"). Also, although primary and secondary exclusion labels are featured in the same category of stylistic labels, secondary labels are still characterized as signaling that the word contains emotional, expressive characterization ("в слове содержится соответствующая эмоциональная, выразительная оценка обозначаемого явления"). One should also note that the vast majority of the labels used in this dictionary are potentially excluding. Those that are not excluding are the following: literary ("книжное"), high register ("высокое"), official/formal ("официальное"), and specialized/technical ("специальное"). In addition, there is the aforementioned colloquial category, which is declaratively non-excluding. The distribution of the relevant labels in this dictionary is presented in Table 4.

Primary exclusion labels

| English | Russian | Number |
| --- | --- | --- |
| regional | обл. – областное | 222 |
| non-standard colloquial | прост. – просторечие, просторечное | 3230 |
| old | стар. – старое | 225 |
| obsolete | устар. – устаревшее | 2604 |
| Total | | 6281 |

Secondary exclusion labels

| English | Russian | Number |
| --- | --- | --- |
| vituperative | бран. – бранное | 51 |
| rude | груб. – грубое | 12 |
| ironic | ирон. – ироническое | 412 |
| disapprobatory | неодобр. – неодобрительное | 784 |
| contemptuous | презр. – презрительное | 128 |
| disparaging | пренебр. – пренебрежительное | 114 |
| facetious | шутл. – шутливое | 652 |
| Total | | 2154 |

Non-exclusion labels

| English | Russian | Number |
| --- | --- | --- |
| high register | высок. – высокое | 766 |
| literary | книжн. – книжное | 1961 |
| official/formal | офиц. – официальное | 267 |
| colloquial | разг. – разговорное | 960 |
| specialized/technical | спец. – специальное | 3636 |
| Total | | 7590 |

Table 4: Exclusion and non-exclusion lables in Ožegov and Švedova (1992)

This dictionary contains 40,229 entries. The total number of labels used is 16,025 (40% of the number of entries), with exclusion labels comprising 21% (primary exclusion 16%, secondary exclusion 5%) while non-exclusion labels account for 19% of all instances of labelling. Given that some entries contain more than one label, the number of labeled entries is somewhat smaller than 40%. These numbers are more interesting in that they show the relationship between exclusion and non-excusive labels (their distribution is approximately the same) as well as between primary and secondary exclusion labels (the former are used four times more frequently than the latter).

## 5.2  Polish Monolingual Dictionary

Żmigródzki et al. (2012) an impressive whopping 1520 words to discussing usage labels (which they call qualifiers, "kwalifikatory"). They also provide a very elaborate classification of these labels. The following types are distinguished:

  – chronological "chronologiczne"
  – frequency "frekwencyjne"
  – stylistic "stylistyczne"
  – related to geographical scope "dotyczące zasięgu geograficznego"
  – related to societal scope "dotyczące zasięgu środowiskowego"
  – specialized "specjalistyczne"
  – expressive "ekspresywne"

Only the category of frequency labels does not contain any exclusion labels. The category "specialized" contains mostly non-excluded items but it also contains excluded professional colloquial usage. The labels used in this dictionary and their frequency is presented in Table 5.

Here again we can see that primary exclusion labels are used much more frequently than their secondary counterparts (nearly five times more). Given that this dictionary deploys an array of specialized labels, it would be difficult to count all non-exclusion labels. However, the fact that the label "literary" (książkowe) has the count of 2111 and "administrative" (urzędowe) has the count of 206, may indicate that this dictionary uses exclusion and non-exclusion labels, just like we found in the Russian dictionary. One can also say that this Polish monolingual dictionary is still a work in progress (it currently contains around 12,000 entries), which may change the ratio of the labels once the project is completed. One noticeable difference in relation to the Russian dictionary is the fact that there is no division between standard colloquial and non-standard colloquial. The label "colloquial" is defined here as units used in informal situations, impossible to use in formal situations ("jednostki używane w sytuacjach nieoficjalnych, niemożliwe do użycia w sytuacji oficjalnej").

Primary exclusion labels

| *English* | *Polish* | *Category* | Number |
|---|---|---|---|
| former | dawne | temporal | 98 |
| colloquial | potoczne | stylistic | 2698 |
| Varsovian | warszawskie | geographical | 2 |
| Cracovian | krakowskie | geographical | 15 |
| Posnanian | poznańskie | geographical | 1 |
| Silesian | śląskie | geographical | 1 |
| Lvovian | lwowski | geographical | 0 |
| Vilniusian | wileńskie | geographical | 1 |
| youth | młodzieżowe | societal | 23 |
| criminal | przestępcze | societal | 8 |
| social environment | środowiskowe | societal | 49 |
| Total | | | 2896 |

Secondary exclusion labels

| *English* | *Polish* | *Category* | Number |
|---|---|---|---|
| pejorative | pejoratywne | expressive | 300 |
| scornful | pogardliwe | expressive | 63 |
| vulgar | wulgarne | expressive | 99 |
| facetious | żartobliwe | expressive | 223 |
| Total | | | 685 |

Table 5: Exclusion lables in Żmigródzki et al. (2012)

## 5.3 Serbian Monolingual Dictionary

Vujanić et al. (2011), a Serbian monolingual dictionary with approximately 80,000 entries, devotes a brief 43-word note stating that some words contain etymological, sphere-of-usage, and subjective-evaluation labels. We can thus see which labels are used from their concrete use and the list of abbreviations. It is obvious that the authors are aware of the different nature of primary exclusion labels (they are called sphere-of-usage labels, which also includes non-exclusion labels) and secondary exclusion labels (called subjective-evaluation labels, which can also contain some non-exclusion labels). Relevant labels are presented in Table 6 along with their frequency.

It is interesting to note here that the ratio between primary and secondary exclusion labels is quite different than in the two dictionaries analyzed previously. Primary exclusion labels are used here just somewhat over twice as much as secondary exclusion labels. Just like in the case of the Polish dictionary, it would be close to impossible to count all non-exclusion labels, but the fact that the label "administrative" (административно) has only 38 uses and literary (књижевно) just 120, most of which are actually literary criticism terms, points to the fact that this dictionary uses non-exclusion labels much less than the previous two dictionaries.

Primary exclusion labels

| *English* | *Serbian* | Number |
|---|---|---|
| archaic | архаично | 199 |
| old literary style | старокњижевно | 11 |
| colloquial | разговорно | 961 |
| regional | регионално | 181 |
| folksy | народски | 708 |
| slang | шатровачки | 23 |
| jargon | жаргонски | 218 |
| school student speech | ђачки | 24 |
| Total | | 2325 |

Secondary exclusion labels

| *English* | *Serbian* | Number |
|---|---|---|
| vulgar | вулгарно | 70 |
| expressive | експресивно | 83 |
| familial/homely | фамилијарно | 51 |
| pejorative | пејоративно | 518 |
| derisive | подругљиво | 22 |
| contemptuous | презриво | 12 |
| ironic | иронично | 240 |
| facetious | шаљиво | 62 |
| Total | | 1058 |

Table 6: Exclusion labels in Vujanić et al. (2011)

## 5.4 Slovenian Monolingual Dictionary

Bajec et al. (2000), a Slovenian monolingual dictionary of 97,669 entries, devotes 774 words in the front matter to discussing relevant types of labels. The authors divide the labels into the following categories a. terminological, b. stylistic-stratal, c. expressive, d. temporal-frequency, and e. special normative. The latter four types of labels are relevant in the process of excluding words and their features from standard Slovenian.

Stylistic-stratal labels comprise some non-exclusion labels (biblical, literary, poetic, administrative journalistic, elevated) but also exclusion labels (dialectal, child-language, colloquial, lower colloquial, slang). The authors divide those into two clear groups, listing the first group of labels under a. and the second under b.

Expressive labels include the following: expressive, euphemistic, ironic, hypocoristic, low, contemptuous, facetious, and vulgar. Although all of them can potentially be excluding, that role is reserved primarily for negatively connoted labels.

Temporal-frequency labels encompass the following items: increasing use, diminishing use, rare, old, and obsolete. The latter two are primary exclusion labels, while the former three remain non-excluding.

Primary exclusion labels

| *English* | *Slovenian* | Number |
|---|---|---|
| Incorrect | nepravilno | 24 |
| Unestablished | neustaljeno | 92 |
| Lower colloquial | nižje pogovorno | 503 |
| Colloquial | pogovorno | 2133 |
| Dialectal | narečno | 1958 |
| Slang | žargonsko | 512 |
| Child-language | otroško | 111 |
| Old | starinsko | 4793 |
| Obsolete | zastarelo | 3893 |
| Total | | 14019 |

Secondary exclusion labels

| *English* | *Slovenian* | Number |
|---|---|---|
| Expressive | ekspresivno | 12355 |
| Low | nizko | 217 |
| Contemptive | slabšalno | 1447 |
| Facetious | šaljivo | 254 |
| Ironic | ironično | 178 |
| Vulgar | vulgarno | 123 |
| Total | | 14574 |

Table 7: Exclusion labels in Bajec et al. (2000)

Finally, a remarkable characteristic of this dictionary, when we compare it with three other Slavic monolingual dictionaries, is the fact that it contains prescriptive labels, i.e., the most direct form of exclusion from standard Slovene. There are two items of this kind: incorrect, and non-established.

Exclusion labels, along with their frequency, are presented in Table 7. As can be seen, owing to the profuse use of the label "expressive", the number of primary and secondary labels is approximately the same.

Again, given that there is no unified "specialized/technical" label, it would be difficult to point out how many non-exclusion labels are used but the fact that "literary" (knjižno) has 9809 uses and journalistic (publicistično) 2356, one could conclude that non-exclusion labels are used in similar magnitude as their exclusion counterparts.

# 6 Conclusions

For our purposes here, we defined the standard language in its narrower sense (see section 2), as those language varieties subject to standardizing efforts by codifiers, linguistic experts, and norm authorities. This would include administrative, legal, scholarly language, news, etc. and exclude informal colloquial and literary language (which are typically not the target of normative interventions). The question ensuing from the review of four Slavic monolingual dictionaries is what strategies the authors of these dictionaries deploy in excluding lexical items or their features from the standard language in its narrower sense. In other words, the question is how these dictionaries serve their users (e.g., language editors) who need to verify if the word or its feature is acceptable in an edited standard language text.

What follows from the review in section 5 is that there is a variety of means by which the exclusion from the standard language can be achieved. The range of strategies goes from a very rare direct prohibition (The Slovene label "incorrect" and the Russian "non-standard colloquial"), pointing to the fields traditionally considered non-standard (e.g., dialect and slang), to pointing to the features incompatible with the standard language as defined here (for example, "vulgar"). We saw that all analyzed dictionaries use primary exclusion labels, pointing to non-standard fields of usage, and secondary exclusion labels, pointing to their non-standard features. Primary exclusion labels are generally more frequent but the number of secondary exclusion labels is substantial. Both categories of labels, primary and secondary, have fuzzy edges. With the exception of the Russian dictionary, where the category of non-standard colloquial is clearly established, in all other dictionaries colloquial actually means informal and the judgements of the acceptability in the standard language (as defined here) may vary. To a lesser degree this is true for old and archaic too. In secondary exclusion labels the categories like familial or facetious bring about the same kind of uncertainty about acceptability. We also saw (demonstrated in the case of the Russian dictionary and confirmed in the other three dictionaries) that the authors are equally concerned with marking certain fields within the standard language (e.g., technical use) as they are with excluding items from the standard. The use of exclusion and non-exclusion labels is equally substantive.

A part of the problem with fuzzy edges of primary and secondary inclusion labels is that, through a historic association of the standard with literary language (in spirit and often in name too), Slavic cultures maintain mostly the broader (cohesive) notion of the standard language. While the distinctive idea remains to be enforced in publications such as manuals of orthography, grammatical and stylistic reference works, dictionaries seem to largely neglect the narrower (distinctive) notion of the standard language. In order to eliminate or at least reduce user's uncertainty as to the acceptability of some items in the standard language, one would need to define the range of the standard language in its narrower sense and deploy the labels accordingly by determining the cohort of primary and secondary labels not compatible with standard language texts.

# References

Ammon, Ulrich. 1995. *Die deutsche Sprache in Deutschland, Österreich und der Schweiz: das Problem der nationalen Varietäten.* Berlin: de Gruyter.

Ammon, Ulrich. 2004. "Standard Variety / Standardvarietät." In *Sociolinguistics / Soziolinguistik. An International Handbook of the Science of Language and Society / Ein internationales Handbuch zur Wissenschaft von Sprache und Gesellschaft*, edited by Ulrich Ammon et al., Vol. 1 (Second edition), 273–283. Berlin: de Gruyter.

Armstrong, Nigel, and Ian E. Mackenzie. 2013. *Standardization, Ideology and Linguistics.* Houndmills: Palgrave Macmillan.

Bajec, Anton et al. 2000. *Slovar slovenskega knjižnega jezika.* Ljubljana: Slovenska akademija znanosti in umetnosti, http://bos.zrc-sazu.si/sskj.html (main body), http://www.fran.si/130/sskj-slovar-slovenskega-knjiznega-jezika/datoteke/SSKJ_Uvod.pdf (front matter)

Béjoint, Henri. 2010. *The Lexicography of English.* Oxford: Oxford University Press.

Davies, Winifred. 2012. "Myths we live and speak by. Ways of imagining and managing language and languages." In *Standard Languages and Multilingualism in European History*, edited by Matthias Hüning et al., 5–69. Amsterdam: John Benjamins.

Dittmar, Norbert. 2004. „Umgangssprache – Nonstandard / Vernacular – Nonstandard." In *Sociolinguistics / Soziolinguistik. An International Handbook of the Science of Language and Society / Ein internationales Handbuch zur Wissenschaft von Sprache und Gesellschaft*, edited by Ulrich Ammon et al., Vol. 1 (Second edition), 250–262. Berlin: de Gruyter.

Fishman, Joshua. 2010. *New Perspectives on Language and Education: European Vernacular Literacy: A Sociolinguistic and Historical Introduction.* Bristol: Channel View Publications.

Geeraerts, Dirk, Gitte Kristiansen, and Yves Peirsman. 2010. *Cognitive Linguistics Research: Advances in Cognitive Sociolinguistics*, Berlin: de Gruyter.

Hartmann, R. R. K., and G. James. 2001. *Dictionary of Lexicography.* London: Routledge.

Hausmann, Franz Joseph. 1989. "Die Markierung im allgemeinen einsprachigen Wörterbuch: eine Übersicht." In *Dictionaries. An International Encyclopedia of Lexicography*, edited by F.-J. Hausmann, O. Reichmann, H. Wiegand, and L. Zgusta, 649–657. Berlin: de Gruyter.

Kroskrity, Paul V. 2010. "Language ideologies – Evolving perspectives." In *Society and Language Use. Handbook of Pragmatics Highlights 7*, edited by Jürgen Jaspers et al., 192–211. Amsterdam: John Benjamin.

Milroy, James, and Lesley Milroy. 2012. *Authority in language: Investigating standard English.* 4th edition. Oxford: Routledge.

Ožegov, S. I., and N. Ju. Švedova. 1992. *Tolkovyj slovar' russkogo jazyka.* Moskva: Az'.

Romaine, Suzanne. 2000. *Language in Society: An Introduction to Sociolinguistics.* 2nd Edition. Oxford: Oxford University Press

Smakman, Dick. 2012. "The definition of the standard language: a survey in seven countries." *International Journal of the Sociology of Language* (218): 25–58.

Svensén, Bo. 2009. *A Handbook of Lexicography. The Theory and Practice of Dictionary-Making.* Cambridge: Cambridge University Press.

Ščerba, L. V. 1957. "Sovremennyj literaturnyj russkij jazyk." *Izabrannye raboty po russkomu jazyku* 110–129.

Takahashi, Hideaki. 2004. "Language Norms / Sprachnorm." In *Sociolinguistics / Soziolinguistik. An International Handbook of the Science of Language and Society / Ein internationales Handbuch*

*zur Wissenschaft von Sprache und Gesellschaft*, edited by Ulrich Ammon et al., Vol. 1 (Second edition), 172–179. Berlin: de Gruyter.

Verkuyl, Henk J., Maarten Janssen, and Frank Jansen. 2003. The codification of usage by labels. In *A Practical Guide to Lexicography*, edited by Piet van Sterkenburg, 297–311. Amsterdam: John Benjamins.

Vujanić, Milica et al. 2011. *Rečnik srpskoga jezika.* Novi Sad: Matica srpska.

Yong, Heming, and Jing Peng. 2007. *Bilingual Lexicography from a Communicative Perspective.* Amsterdam: John Benjamins.

Zgusta, Ladislav et al. 1971. *Manual of Lexicography.* The Hague: Mouton.

Żmigrodzki, Piotr et al. 2012. *Wielki słownik języka polskiego*, Kraków: Instytut Języka Polskiego PAN. http://wsjp.pl.