

Food as Text

UDC 811.163.41'322.2 811.163.4'374

DOI 10.18485/infodhca.2019.19.2.7

ABSTRACT: This paper aims to describe the initial steps in building a diachronic and interdisciplinary corpus that would present different aspects of the language of food in the Serbian language. This corpus consists of culinary recipes, ethnographic and anthropological studies and other different testimonies about food in Serbia dating from the second half of the 19th century onward. Such corpus offers insights into eating habits in Serbia and how they changed under different cultural influences. The problems of automatic processing of this corpus are discussed, some experiments that rely on the use of existing lexical resources are presented, and directions for future work are given.

KEYWORDS: domain corpus, culinary, food, computing, electronic dictionary.

PAPER SUBMITTED: 19 December 2019

PAPER ACCEPTED: 25 December 2019

Duško Vitas

vitas@matf.bg.ac.rs

*University of Belgrade
Faculty of Mathematics
Belgrade, Serbia*

Translated from Serbian:

Ana Popović

1 Introduction

This paper aims to describe the initial steps in building a corpus that would present different aspects of the language of food in the Serbian language. Food is talked about from various perspectives, from the daily preparation of meals to the consideration of the nature of food from the point of view of a number of scientific disciplines or gastronomic accounts. This language, due to its everyday use, seems at first sight quite ordinary. Yet when considered over an extended period of time, it is possible to identify complex layers of different cultural and socio-economic influences to which it is a testament.

The fact that the language of food has been neglected is illustrated, among other examples, by the fact that in the Dictionary of the Serbian Academy of Sciences and Arts (RSJ, 2011), in the twenty volumes published

to date, only about fifty entries have the usage label *kuv.* meaning “kuvarski termin, kulinarstvo” (“culinary term, cooking”). In the dictionary’s last published volume, for instance, the entry for the word *pasulj* (*beans*) is labelled as a botanical and agronomic term, but not a culinary one. The definition of the primary meaning says “...leguminous plant... used in nutrition as a vegetable, the seed of this plant, a dish made with this plant, *grah*”. It should be noted that there are two distinct meanings here: *leguminous plant* and its *seed* are botanical and agronomic concepts, while a *dish* made with this plant is a culinary concept. Moreover, most examples given for the primary meaning refer to the dish,¹ and only one each to its botanical and agronomic meanings. As further testimony to the complexity of the term, I will cite a personal experience – namely, I recently attended a three-hour scholarly discussion between biologists, chemists, physicochemists, doctors and agronomists on the properties of leguminous plants, whose most important representative is the *bean*. Only once did the discussion on the complex biochemical properties of this plant family touch on *beans* as a dish (in the form of *prebranac*).

Traditional lexicographical studies of the language of food in Serbian are based either on ethnolinguistic studies or on the excerption of culinary terms primarily from existing dictionaries. Of the extensive body of literature on the subject, we will cite here only the most recent. One of the ethnolinguistic works describing the culinary vocabulary of a particular geographical region in great detail, but without a replicable and precise specification of the corpus used in the study, is (Mirilov, 2016). A different approach has been used in (Radonjić, 2016), where culinary vocabulary is described based on descriptions in the Matica Srpska Dictionary (RSJ, 2011), also drawing on other dictionaries as well as cookbooks. As the relevant material does not include appropriate labels, the excerption had to be done “by hand” (cf. footnote 13) and, inevitably, partially. Both approaches deal with the “ordinary” names of dishes, although the concept of “ordinary” or common is very hard to determine without an adequate referential corpus.

In addition to lexicographical and ethnolinguistic sources about the language of food, numerous other sources describe or testify to dietary habits from different perspectives – ethnological, anthropological, historical, sociological, philosophical... Although viewed from different perspectives, the

¹ A search of the Corpus of Contemporary Serbian Language confirms that the entry refers primarily to the dish rather than the plant (<http://www.korpus.matf.bg.ac.rs/korpus/>).

described object remains the same – foodstuffs, dishes and ways of preparing them. The paper aims to explore the possibilities of forming an electronic corpus that would present these various views of food and enable not only a lexical search, but also a search that would include historical and geographical information on the origins of the vocabulary used, and also about its evolution. The fluctuations in the occurrence of the word *bulgur* can serve as an example. It appears in (Тројановић, 1983),² where it says that “in our country...*bungur*³ is also highly prized”, but it does not occur even once in Midžina’s cookbook (Поповић Миџина, 1878). About half a century later, in (Марковић, 1959) there is no mention of *bulgur*, but (Zirojević, 2019) provides a detailed description of *bulgur* in a separate section entitled “Three (not) forgotten foods”, as the food had become fashionable again, due to the influence of medical research.⁴

Such a corpus could help synthesise information scattered across different papers, dictionaries, studies and cookbooks, information that is not only linguistic and lexicographical, but also belongs to the domain of the humanities, e.g. ethnology, anthropology or history. Thus, the corpus could provide the basis for an encyclopedic study of the language of food in the Serbian language.

In addition to the collection of material, there is the question of instruments for its processing. On the one hand, these are lexical resources enabling the adequate indexing of the corpus, and on the other, a complex corpus administration system which, apart from standard functions, must also establish connections not only within the corpus material but also with external sources such as geographical maps or encyclopedic sources. Moreover, preliminary processing should include text normalization, in terms both of vocabulary and of other markers used (e.g. the system of measurement).

A part of these resources for Serbian have already been constructed over recent years through the development of specific dictionaries and grammars for the Unitex system⁵ (Крстев and Лазвић, 2015) and additions to culinary terms of the WordNet type semantics network for the Serbian language (Вујићић Станковић and Пајић, 2014). Part of the necessary material – a corpus of cooking recipes collected from the web – is described in (Vujičić-Stanković et al., 2014). This corpus gives partial insight into the

² The study was first published in 1898.

³ A dialect variant of *bulgur*

⁴ In the Corpus of the Contemporary Serbian Language, *bulgur* occurs ten times (all ten occurrences are after 2008).

⁵ [Grammar-based corpus processing suite](#)

contemporary cooking vocabulary, but it lacks certain information, notably on the socio-cultural (e.g. rural – urban), geographical and historical origin of recipes. A more complex organization of culinary material can be found in (Stošić et al., 2017), where different culinary sources are organised as a multimedia document.⁶

2 The structure of the corpus

The current version of the language of food corpus comprises different types of written sources on culinary topics. In other words, the corpus consists of the following types of texts:

- terminological monolingual and multilingual dictionaries and glossaries (for instance, (Baničević and Popović, 2010), (Vukov, 1954));
- cookbooks in Serbian (such as (Поповић Мицина, 1878) or (Марковић, 1959)) or translated into Serbian, e.g. (Пелапрат, 1973);
- doctoral dissertations addressing culinary topics, e.g. (Mirilov, 2016);
- Serbian-language recipe collections collected from the web, including user comments if available;
- culturological, ethnographic and anthropological studies (such as (Montanari, 2011), (Тројановић, 1983), (Zirojević, 2019), (Радуловачки, 1996) and (Милорадовић, 2014));
- newspaper articles from general sources or specialised food magazines;
- a collection of menus (restaurant menus or brochures);
- monographs on food (e.g. (Мијо, 2012) or (Onfre, 2002));
- historical accounts (e.g. (Фотић, 2005));
- excerpts from literary works in which food is mentioned (e.g. excerpts from (Петроније, 1976), (Игњатовић, 1949) or (Балзак, 1933)), and
- textbooks (e.g. (Портић, 2011))

This list provides insight into the heterogeneous character of the corpus, because the nature and inner organization of the sources described varies widely. Apart from being different in character, the texts are not equally accessible: certain types of sources described are not archived and are impossible to find. A case in point are restaurant menus testifying to changes in urban dietary habits:⁷ a valid and temporally well spaced sample of such

⁶ A students' multimedia document "Back Then Eating Was Good"

⁷ (Голубовић, 2007) provides an exhaustive list of Belgrade hotels, restaurants and inns from the mid-19th century, but with hardly any mention of the dishes they served. One of the few descriptions cites an excerpt from Miloš Crnjanski's *Belgrade* (pp. 46-47).

documents is not available to date. The importance of these marginalised documents is explained in (Витас, 2018): the habitual diet of a certain urban population cannot be inferred from cookbooks. On the other hand, the language of food varies depending on the type of source. Descriptions range from the “algorithmicised” text of recipes or the simple language of cookbooks to sophisticated consideration of flavours or complex cultural-historical influences the result of which is a certain dish. This stylistic heterogeneity also raises the question of weighting, which would enable the valorization of certain types of views of the corpus.

The structure of the documents in the corpus is also heterogeneous. Some texts have a normal, ordinary structure, while others require specific tagging. For one thing, the texts of recipes are non-uniform. Contemporary recipes, collected from recipe websites, tend to have a precise structure comprising the name of the dish, important ingredients, required quantities, time necessary for the preparation of the dish, etc., followed by the preparation procedure (sometimes broken down into steps). Older recipes, on the other hand, tend to state the name of the dish and the preparation procedure which includes those parts that are stated separately in contemporary recipes. In Section 3, we discuss the possibility of automatically translating the old into the new structure. Menus are also specifically structured in the form of lists or lists of lists (when ingredients are listed).

In addition to the normalization of specifically structured documents, there is the issue of the lexical normalization of the corpus. Namely, lexical variations stem both from the distribution of regional or archaic words, and from the nonstandard way in which new culinary terms are adopted from foreign sources. This problem has been partly solved in (Vujičić-Stanković et al., 2014). Examples include the word *ožica* in (Тројановић, 1983) for *kašika* (spoon), or *prevrtača* in (Ердељановић, 1908) for *palačinka* (crepe/pancake). In (Поповић Миџина, 1878) we find that different ways of *pohovanje* (frying in breadcrumbs/batter) are reduced to *prženje* (frying) (e.g. *pržene teleće nožice*), or that *musaka* referred to in (Ердељановић, 1908) is *modri patlidžan za 6 osoba* (aubergine for six persons). The tracing of terminological evolution in a corpus thus conceived means that cases such as these will also have to be included in the search.

Finally, in such heterogeneous sources, sections relevant for the corpus have been presented in different ways. Unlike the class of recipes, which give a straightforward description of a dish, in different types of studies recipes or language about the use of a particular food is inserted into a wider context that is not necessarily relevant. The separation of parts that constitute a

recipe from the parts that describe the processing of foods or flavours would have to be the subject of separate tagging which would need to be described.

3 Computational issues of processing cooking texts

A corpus conceived in such a way is not only important for ethnographic and lexicographical research, but also offers interesting material for different computational experiments which we will briefly describe here. These experiments would potentially lead to the defining of a formalised language of cooking, whose immediate application would be in programmable cooking robots,⁸ and in query systems intended for assisting in food preparation.⁹ Some of the tasks are applications that, apart from being standalone applications, could be used in the normalization and tagging of corpus texts.

Let us consider some of these tasks that would have to be resolved for the Serbian language in order for these applications to be built. One group of tasks involves supplementing cooking resources already developed for the Serbian language, and another regards systems for analyzing and transforming corpus texts.

The first group of required resources includes supplementing constructed cooking dictionaries which are partly described in (Vujičić-Stanković et al., 2014), (Stijević et al., 2016). Specific regional, archaic and similar vocabulary has to be processed and added to the system of electronic dictionaries so that the corpus can be properly tagged. The vocabulary necessary for recognizing specific named cooking entities occurs here as a separate subsystem. Approximate measurements are analysed and discussed in (Krstev et al., 2014), but, in addition to these, there are hidden entities in the texts whose value has to be inferred (or checked) indirectly. Examples include information on preparation time, temperature etc., but also a group of names of ingredients, preparation procedures and dish names that include proper nouns in their basic or derived form (Krstev et al., 2019), (Vujičić Stanković and Pajić, 2015).

Semantic relationships, which are partly built into the system of electronic dictionaries for the Serbian language and also into the Serbian WordNet, would have to be significantly expanded with qualifiers such as “re-

⁸ Cooking robots available today are capable of preparing dishes for which the procedure has been previously defined, but the user cannot instruct them to prepare his/her own recipe (Vitas and Krstev, 2016).

⁹ Such a system was proposed through IBM’s *Chef Watson* system

gional”, “archaic”, etc., as well as data on energy value, allergens etc. including references to corresponding encyclopedic content. It should be noted that words of similar meaning can be used synonymously, but can also refer to a substitute for a certain ingredient, which represents a separate relationship within the semantic network. For example, *puter*, *buter*, *maslo* and *maslac* are part of the same WordNet synset,¹⁰ but *margarin* belongs to a different synset. Due to their interchangeability, a separate relationship is needed to indicate possible ingredient substitution.

Multi-word units in the language of cooking, in addition to their usual meaning, can also be the name of a new culinary concept, particularly in terms of names of dishes. Their structure both at the level of ingredients and cooking utensils, and that of dishes, has not been the subject of separate lexicographical processing in Serbia. In the case of culinary innovations, which usually represent “unusual” dishes, the entire description of ingredients can appear as a lexical unit, for instance, *pohovani bri sa kulijem od šumskog voća* (breaded fried brie with forest fruit coolie) or *pate od čvaraka sa musom od kozijeg sira* (pork rilette with goat cheese mousse). This phenomenon of a lack of a “standard” lexeme for a certain dish occurs in other languages as well, as shown by analyses provided in (Gerhardt et al., 2013).

An extremely complex lexicographical issue is the question of multilinguality. Namely, WordNet, through a system of interlingual indices, enables the pairing of equivalent concepts. But the question of cooking recipes in a certain culture is beyond the possibilities of equivalence description at the synset level. It is a specific transfer from one language to another that depends to a great extent on the local (or national) system of food concepts. For example, *musaka*¹¹ is a Balkan concept that exists in all European languages, but the local manner of preparation varies to such an extent that it is not always certain that it refers to the same type of dish. On the other hand, there are conceptually similar dishes prepared with somewhat different ingredients outside the Balkan region (e.g. Italian *lasagna* or French *hachis parmentier*). This raises the question of identification of dishes where the preparation procedure is very similar although both the ingredients and the names of dishes vary between languages. One solution is connecting the

¹⁰ A synset groups together words with similar meanings into a semantic network graph node. This node is connected to other synsets by semantic relations.

¹¹ In (Courtine, 1986) *musaka* is defined as a Balkan dish of alternating layers of aubergine slices and minced lamb, usually with a coating white sauce. Unlike this simplified presentation, (Марковић, 1959) gives more than 30 different recipes for this dish.

name of the dish to the family of recipes it is represented by, and then establishing equivalence between similar procedures in different languages. Such a procedure could lead to the establishment of an abstract schema of dishes that would enable lexical transfer from one language into another.¹² Nevertheless, a specific restriction on the internationalization of cuisine, despite culinary globalization, are local ingredient availability as well as local culinary techniques and gastronomic habits. As an example, *baget(a)* can be bought in most Belgrade bakeries but, apart from its name, it has little in common with the French *baguette*. Another issue is establishing a connection, whenever possible, between a certain dish and its historical and geographical origins. In (Поповић Миџина, 1878), for instance, we can find examples such as *teleće pržoljice* (*Kalberne Schnitzel*, *teleće šnicle*), this being the only occurrence of the word *šnicla* in the cookbook. The word *pržoljica*, however, has a broader meaning, because there are also *ovnujske pržoljice* – *Cotelette* – *rebarca*. In this case we can see the German and French influences on the forming of contemporary culinary names (*šnicla*, *kotlet*). Interestingly, none of the three names (*pržoljica*, *šnicla*, *kotlet*) occur in either (Тројановић, 1983) or (Ердељановић, 1908), while (Марковић, 1959) lacks only *pržoljica*: *šnicla* and *kotlet* had probably displaced *pržoljica* as a result of culinary refinement.

In addition to these requirements in the description of lexical structures and relations in culinary language, it is possible (partly starting from existing resources) to develop applications necessary for corpus normalization. Among such applications are programs that analyse possible combinations of ingredients, taking into account the frequency of their co-occurrence. For example, *breskva* (peach) and *ananas* (pineapple) are only very rarely connected to *so* (salt). A similar application would be one that analysed similar (or identical) recipes where the same name can involve different ingredients and procedures and vice versa, where different names are used for the same dish made from identical (or similar) ingredients and prepared using the same procedures.

Useful applications that would improve existing searchable recipe collections are numerous, and we will mention them only as an idea for some interesting computational experiments:

¹² Conceptual equivalence of two dish names in different languages might be established by identifying a similar procedure using different ingredients.

- an application that determines the level of difficulty of preparing a certain recipe based on an analysis of the operator (cooking verbs) and ingredients;
- an application that determines the level of difficulty of preparing a certain recipe based on an analysis of the operator (cooking verbs) and ingredients;
- an application that determines the appropriate season for preparing a recipe based on the seasonal availability of ingredients;
- an application that automatically determines the course (starter, main course...);
- an application that identifies recipes appropriate for a certain diet or the risks of consuming a certain dish or ingredient (e.g. tolerance of pungency on the Scoville scale is subject to individual and regional differences).

4 On the processing of the corpus and initial results

A corpus thus conceived transcends the possibilities of the usual ways of processing in view of the complex structure of heterogeneous documents and the nature of possible queries. In other words, a search that, in addition to key words, involves other resources – dictionaries and semantic networks, metadata and different inferred or external data – necessitates a system that, in addition to a corpus processing function, enables other types of analysis. In the experimental phase of constructing such a corpus, a software solution could involve the application of the Unitex (Paumier, 2016) system in the stage of corpus preprocessing and tagging, and a system such as TXM¹³ for storing and searching texts (Jačimović, 2019).

We will limit ourselves here to the initial stage of preparation: the normalization and tagging of texts for the future corpus. This stage involves the intensive use of the above-mentioned resources developed for Unitex. Most collected texts have been scanned and read (by means of OCR), using the system described in (Krstev and Stanković, 2019) (Krstev, švalje), and they have been semi-automatically corrected including minimal TEI-tagging (in accordance with the requirements of the EITec project).¹⁴

As part of a preliminary analysis of possible applications of the corpus, we will examine relations between two groups of texts, one that looks at the

¹³ The TXM package, intended for textometry, in addition to a corpus processing system (IMS CQP) also integrates other means of analyzing textual data, such as the R system.

¹⁴ Distant Reading for European Literary History (COST Action CA16204)

phenomenon of food from the perspective of ethnology and anthropology, and another consisting of cookbooks and testifying to culinary practices in a certain period. The texts selected for the experiment are listed in Appendix 6, and hereafter we will refer to them using the letters and numbers assigned to them in the Appendix. These texts cover the period from the late 19th century to the early 21st century, and they provide accounts of dietary habits of the period when they originated and of the environment they describe. Texts A1-A4 describe dietary habits in rural areas, while others testify to urban culinary practices. The original year of publication, and basic data about the size of individual texts in terms of number of words, the number of different words and unrecognised words (as analysed by an electronic dictionary system) are given in Table 1).

| | A1 | A2 | A3 | A4 | A5 | A6 | B1 | B2 | B3 |
|------------|--------|--------|--------|--------|--------|--------|---------|-------|---------|
| year | 1896 | 1908 | 1908 | 1996 | 2014 | 2018 | 1878 | 1915 | 1959 |
| total | 33,913 | 19,266 | 13,899 | 24,914 | 74,633 | 43,189 | 121,531 | 83,08 | 399,358 |
| different | 9,220 | 4,477 | 3,302 | 7,964 | 18,648 | 12,032 | 9,367 | 1,709 | 16,542 |
| repetition | 3.68 | 4.3 | 4.21 | 3.12 | 4.00 | 3.56 | 12.98 | 4.86 | 24.14 |
| err | 956 | 338 | 480 | 442 | 479 | 1,189 | 1,107 | 79 | 175 |

Table 1. Quantitative data on sources

The relationship between the total number of occurrences of simple words and of different simple words shows that the level of repetition in cookbooks is high, which indicates that the basis of the language of cooking – ingredients and procedures – is very limited. The number of unrecognised words in the err category is extremely low for B2 and B3 because these are texts that have been thoroughly processed, and the vocabulary entered into the electronic dictionary. The character of other unrecognised words varies: they come partly from foreign languages (in transcribed or non-transcribed form), but mostly from specific local vocabulary, primarily in the names of dishes, and to a lesser extent from non-standard variations in the names of ingredients, procedures or dishes. Table 2 shows the most common unrecognised words or frequent forms that have not been included in the appropriate inflectional

class of the entry, and their frequency. In A1, for instance, in this stratum we come across the word *supraška* (meaning *vreo pepeo* [hot ashes]) in different forms 18 times, but also the sibilised stem *surutci* eight times (while other forms are included in the pattern <*surutka*>).

Looking at the unrecognised words in Table 2, we notice oppositions in terms of rural-urban, archaic-modern and also regional differences. On the other hand, we note that the context of the unrecognised words, if they occur in multiple sources, indicates different meanings. Thus *vodnjika* in A1 refers to an old drink, like *jabukovača* [apple brandy], so common, according to the author, that it does not need a description; in A2 it is made from pears, and in A4 it is the same as *jabukovača*; it does not occur in other sources. This confirms that along with the names of dishes (and drinks) it is also necessary to state the procedure, because different products are labelled with the same name. Also, linking these sources to the literary part of the corpus makes it possible to interpret the literary content. For example, in Rastko Petrović's novel "Burleska gospodina Peruna, boga groma" we come across both *vodnjika* and *supraška* from A1.

Let us now look at the distribution of some culinary tags in electronic dictionaries described in (Крстев and Лазич, 2015). Let us take as an example the occurrence of plant names (as foodstuffs and as dishes) in certain sources tagged in the e-dictionary as *Bot* and *Food*.

In A1 grains predominate (*žito* (36) [cereal], *pšenica* (24) [wheat], *ječam* (21) [barley], *ovas* (15) [oat], *raž* (7) [rye], *proso* (6) [panicgrass], *krupnik* (5) [spelt], ...), *kukuruz* (17) [corn] and *bundeve* (20) [pumpkin], adding up to over half of the total number of words thus tagged. Fruit and vegetables, such as *kupus* (9) [cabbage], *luk* (8) [onion], *pasulj* (5) [bean], *pirinač* (3) [rice], *krompir* [potato] (3),¹⁵ *jabuke* (3) [apple] i *šljive* [plum] (2), are rarely mentioned. This disregard for vegetables in A1 probably stems from the author's aim to describe the simplest, most traditional diet that could still be attested in the late 19th century.¹⁶ Unlike in A1, in A2 vegetables take precedence over grains: in the text we find *luk* (*beli* [garlic] and *crni* [onion]) (71), *paprike* (45) [pepper], *kupus* (38) [cabbage], *krompir* (23) [potato], *pasulj* (22) [bean], *zelje* (18) [garden patience], *tikve* (but not *bundeve*) (18), *boranija* (11) [green bean], *patlidžan* (*crveni* and *zeleni* = *paradajz* [tomato]) (11), *pečurke* and *gljive* (10) [mushroom], *krastavac* (10) [cucumber], *pirinač*

¹⁵ A1 states that in the mid-nineteenth century potato is cultivate "rarely, and more from curiosity than necessity".

¹⁶ "For these reasons, we shall describe more thoroughly the simple foods of our people, and thus also help westerners to contemplate their past through us."

| A1 | A2 | A3 | A4 | A5 | A6 | B1 | B2 | B3 |
|------------------|---------------------------|----------------------------------|----------------------|---------------------------|--------------------------|------------------------------|---------------------------|-----------------------------|
| supraška (18) | kalenica (28) | (is na pre u) križati (49) | komlov (6) | pomander (28) | dirhem (10) | (za u 0) prigati (239) | (za u) prigati (40) | "A la..." (6) |
| kačkaval (17) | prazi luk (16) | (po is) prigati (46) | trgančiči (5) | fish and chips (25) | kabak (8) | rem (118) | zapraska (15) | drinks (5) |
| jagurt (11) | izmeljati (14) | bungur (12) | uslatko (4) | fast-food (9) | fasulj (7) | zemičkin (104) | rem (10) | (lon šort) drinks (4) |
| varica (9) | gruvanica (8) | nešesta (12) | sacurica (4) | garum (7) | patlidžajin (6) | buavan (90) | gris (9) | prevru (2) |
| kanavac (8) | (na is) križati (8) | ovlaš(e) (11) | švargl (3) | trufa (7) | obiber- čufter (6) | parme- isati (81) | zanski (9) | kvascom (2) |
| surutci (8) | kupusnik (5) | mohuna (10) | ukiselo (3) | princes- krofna (4) | pasuljica (6) | morunin (46) | obiber- sati (6) | vips (2) |
| bungur (7) | vodnjika (5) | češkek (9) | vani- krancle (3) | taan (3) | piper (6) | brkljača (41) | šerpenjica (4) | češnjom (2) |

Table 2. The most frequently unrecognised words in the samples

| | A1 | A2 | A3 | A4 | A5 | A6 | B1 | B2 | B3 |
|--------------|--------|--------|--------|--------|--------|--------|---------|-------|---------|
| total | 33,913 | 19,266 | 13,899 | 24,914 | 74,633 | 43,189 | 121,531 | 8,308 | 399,358 |
| <Bot+Food> | 335 | 438 | 307 | 535 | 1466 | 1299 | 3252 | 263 | 13736 |
| % | 1% | 2.3% | 2.2% | 2.1% | 2% | 3% | 2.68% | 3.2% | 3.5% |

Table 3. The share of plant foodstuffs in the samples

(10) [rice]... and the following fruit: *šljive* (25) [plum], *grožđe* (10) [grape], *jabuke* (8) [apple], in addition to a number of other fruit and vegetables with lower frequencies. In A3, although it was published in the same period as A2, there are several notable differences. Here also *luk* [onion] occurs with the highest frequency (54), followed by *patlidžan* (*crveni* and *modri* [tomato and aubergine]) (23), *kupus* [cabbage] (20), *pšenica* (18) [wheat], *zelje* (16) [garden patience], *pirinač* (13) and *oriz* (11) [rice], *bungur* (12) [bulgur], *kopriva* (10) [nettle], *krompir* (10) [potato], *tikve* (10) [pumpkin], and then *grah* [bean], *mohuna* (= *zeleni grah*) [bean] and *buranija* (together 24) [green bean], *bamje* (6) [okra], *paprika* (4) [pepper] and other less frequently mentioned vegetables. More frequent fruit are *orah* (7) [walnut], *grožđe* (6) [grape] and *jabuka* (5) [apple]. In A4, published a century after the previous sources, and in another region, a change in the frequency of plant foodstuffs is noticeable. Here, *krompir* [potato] has the highest frequency (56), followed by *luk* (39) [onion], *kupus* (31) [cabbage], *paprika* (31) [pepper], *pasulj* (30) [bean], *žito* (27) [cereal], *bundeve* and *tikve* (which are different, taken together 21 [pumpkin and squash]), *paradajz* (17) [tomato], *kukuruz* (16) [corn], *tikvice* (9) [zucchini], *zelje* (7) [garden patience], *boranija* (6) [green bean]... while the most frequent fruit are *jabuke* (30) [apple], *grožđe* (29) [grape], *višnje* (23) [sour cherry], *šljive* (20) [plum], etc.

In contrast to these ethnographic descriptions of plant foods in rural areas, in the anthropological study A5, as a result of the internationalization of urban cuisine, other plants also appear, displacing to a greater or lesser extent those used in rural areas. Thus, apart from *krompir* (77) [potato], we can find *paradajz* (76) [tomato], *bosiljak* (65) [basil], *špargle* (59) [asparagus] and *masline* (57) [olive] as a direct association with Mediterranean cuisine, or *bob* (62) [broad bean], *avokado* (52) [avocado], *jagode* (53) [strawberry], and *kesten* (49) [chestnut] as a testament to new sophisticated or discriminating

tastes. Plants mentioned in the previous sources (A1-A4) have been pushed into the background: *paprika* (53) [pepper], *kukuruz* (39) [corn], *patlidžan* (37) [aubergine], *kupus* (30) [cabbage], *luk* (30 generally as *beli* or *crveni*) [onion and garlic]... while *pasulj* [bean] occurs only 9, and *bundeva* [pumpkin], *tikvice* [zucchini] and *zelje* [garden patience] only 2-3 times. In A6, which gives an account of the history of foods, most frequent are *krompir* (203) [potato], *kukuruz* (155) [corn] and *paprika* (126) [pepper] although they are all imports from Central and South America. Among the most frequent foods are also *pasulj* [bean], in various forms (*grah*, *boranija*, *buranija* – 205 times in total), *patlidžan* (94) [aubergine], *tikva* and *bundeva* (82) [pumpkin and squash], *paradajz* (71) [tomato], *luk* (62) [onion], *heljda* (62) [buckwheat], *spanać* (59) [spinach], *pirinač* (26) [rice].

In cookbooks, B1-B3, the plant base comprises *luk* (*beli* [garlic] and *crni* [onion]), *krompir* [potato], *prininač* [rice], *pasulj* [bean] and *kupus* [cabbage], while the percentage of *paprika* [pepper] and *paradajz* [tomato]¹⁷ has changed significantly between B1 and B3, as shown in Table 4.

| | luk | krompir | pirinač | kupus | paprika | pasulj | patlidžan | paradajz |
|-----------|------------|----------------|----------------|--------------|----------------|---------------|------------------|-----------------|
| | onion | potato | rice | cabbage | pepper | bean | aubergine | tomato |
| B1 | 328 | 160 | 111 | 93 | 65 | 61 | 46 | 0 |
| B2 | 47 | 31 | 24 | 29 | 28 | 31 | 31 | 3 |
| B3 | 1901 | 728 | 437 | 442 | 61 | 203 | 198 | 604 |

Table 4. The most frequent plant ingredients in cookbooks

The listed frequencies actually indicate the ratio of foods within a certain source. Thus *krompir* [potato], almost entirely neglected in A1, has greater frequency than any other plant foodstuff in A5 and A6, and this is also corroborated by the most recent sources (Витас, 2018). In cookbooks, which testify to urban cuisine, potatoes are already – besides onions¹⁸ – the most

¹⁷ *Paradajz* [tomato] is mentioned in B1 only as *crveni patlidžan*, while in B2 both names can be found.

¹⁸ According to (Вивијен, 2008), “the onion,... which is usually served raw, is the plague of Serbian cuisine (p. 179).

important plant food in B1 in the late 19th century. It is interesting to note the evolution of names as a result of observed differences between foods. In A2 *patlidžan* always refers to *paradajz* [tomato], while in B5 it is the generic name for *modri patlidžan* [aubergine] and *paradajz* [tomato]. In A5 it is used in its present meaning (aubergine), while in A6 it is distinguished by colour: green, blue, black and red. It is also uncertain what *bundeva* and *tikva* refer to in certain authors, that is, whether they denote the same plant [pumpkin and squash]. These examples suggest that the names of foodstuffs, just like the names of dishes, are both subject to change over time and geographically determined, and sometimes insufficiently defined.

In the preparatory phase, relying on e-dictionary tags, it is possible to tag texts using morphological grammars in Unitex. One such grammar is shown in Figure 1.

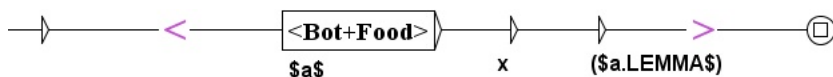


Figure 1. Morphological grammar

This finite-state automaton finds in a text the forms of words tagged as Bot and Food in the dictionary, places them in the `a` variable, and then adds to each form the name of the grammar by which it has been recognised (BOT), the lemma (`$a.CODE.LEMMA$`) and the semantic codes in the entry's field of syntactic-semantic properties (`$a.CODE.SEM$`). An extract of the concordances of different plant names in A6, generated by the grammar, is given in Figure 2.

Similar grammars have been formed for other features in the dictionary such as tags for courses (Course), drinks (Drink), ingredients (Ing), meals (Meal), approximate measures (MesApp) and utensils (Uten). These grammars have been collected into the TAG automaton (Figure 3) which then tags in the text those words that have any of these features.

An example of the application of these grammars to the A6 sample (p. 154) gives the following tagged sequence:

Tu se daje i recept, pa tako saznajemo još da se boranija (boranija, Bot+...+Food) tada pripremala od spanaća (spanać, Bot+...+Food+DOM=Culinary), pirinča (pirinač, Bot+...+Food),

re svega, jelima s kukuruzom, pasuljem, [amarantom](#) i žalfijom. (S) I danas je kukuruz u Meksiku g mu se dodaju razni začini (nana, čubar, [anis](#), mirođija, lovor) koji olakšavaju varenje. (S) Uz t lenim mahunama (lubija sabz), s kiselim [artičokama](#) (kangar) i s vrganjima (karš), uz obavezni d amošnjih oblika i značenja: artičočina, [artičok](#), antričok, antričok, ratičokovina, ratičok, art ričok, antričok, ratičokovina, ratičok, [artišoka](#), artičouka, artičoha i dr. (S)Nalazio se na me), bundevama, paprikama, paradajzom, sa [avokadom](#) i mesom kojim su raspolagali (uključujući glod iška ljuta turšijara, paradajz paprika, [babura](#), kurtočka, somborka, slonova surla... (S)Za ljut k od krasta i od uhobolje, dok je goraki [badem](#) sa šećerom lek za one koji pljuju krv. (S) Živi jo d crvenih i crnih patlidžana, paprike i [bamije](#). (S) Crni i crveni patlidžani ulazili su, uz luk raši i tako dalje (S) Dodaje se, obično, [bamiji](#), pasulju, boraniji, patlidžanu, krompiru, kupusu crveni patlidžani ulazili su, uz luk i [bamiju](#), i u posnu papazjaniju. (S) A poput paprike, i fr u, pa čak i ispod palmi i kraj plantaža [banana](#). (S) I dok u razvijenim zemljama njegova proizvod dobro i dosta crna luka, i sa glavicom [bela luka](#) i dve tri ljute paprike, mete u lonac da se s korijandera, kumin, aleva paprika, so, [hiber](#) (uz druge začine, brašno i prašak za pecivo). (S) kompleks „koji čine kukuruz, pasulj (ili [bob](#)) i tikvice (ili bundeve); ove poslednje leže u podn ture koja počinje da zamenjuje zatečeni [bob](#), od čijih su se zrna gotovile „lepe posne pititije“. lim mlekom. (S) Za Vuka Karadžića (1818) [boranija](#) je „die noch grtinen Fisolen (lat. phaseoli vi) U Sremu, uz pekmez od šljiva, jabuka, [bresaka](#), ringlova, drenjina, <!-- p n = 144 --> grožđa, ečenka, kuhanica - pominju se i strana: [bundeve](#), dulek, jurget, kakab, misirača. (S) Ime ove pos jne, patlidžarnik, pa pečeni sus luk, i [buranija](#). (S) A ova poslednja spravljala se tako što se zovu na čeruiš. (S) Pripremaju se kao i [buranija](#) od tikve.“ <!-- p n = 155 --> I danas u Bosni kuhinji bilo je sedam jela u kategoriji [buranije](#): sa spanaćem (esfenadž), s tikvicama (kadu), s buranija. (S) A priča počinje upravo sa [buranijom](#), tačnije, sa Buran bint Hasan ibn Sahl, ženom i sam navodi tri: buraniju od tikvica, [buraniju](#) od mahuna i buraniju od poriluka. (S) Pišući u ističe Alija Lakišić, i sam navodi tri: [buraniju](#) od tikvica, buraniju od mahuna i buraniju od p an, na salatu, a ponajviše u tepsiju na [buraniju](#), ili kako mnogi zovu na čeruiš. (S) Pripremaju raniju od tikvica, buraniju od mahuna i [buraniju](#) od poriluka. (S) Pišući u Tradicionalnoj kuhinji u Bosni Luka Grdić Bjelokosić pominje i [buraniju](#). „Pravi se ponajviše od tikve, a može se način ak luka, paradajza, gljiva, šargarepe i [celera](#), naziva našim imenom zakuska. (S) Budući da u pot o jelo pasuljica (sa šećerom, orasima i [cimetom](#)). (S) Ova mahunarka koristi se u mnogim lokalnim kuhinji. (S) Može se koristiti zajedno s [cimetom](#), kardamonom, korijanderom, karanfilićem, bibero edoniji ovako: „Usitni se dobro i dosta [crna luka](#), i sa glavicom bela luka i dve tri ljute papr zma. (S) Budući da utiče na metabolizam, [čaj](#) od osušenih mahuna treba da ubrzava i mršavljenje. (oga svrstava u grupu vitamina R Otuda i [čaj](#) od heljde, zbog svog visokog sadržaja rutina, utiče bliku čaja za ublažavanje kašlja, kao i [čaj](#) od cvetova i lišća kod arterioskleroze, a svežim li kamena, a zaustavlja i krvarenje. (S) Uz [čaj](#) od kukuruze svile, koristi se i njen tečni ekstrak koristili su se cvetovi heljde u obliku [čaja](#) za ublažavanje kašlja, kao i čaj od cvetova i lišč sno da mu se dodaju razni začini (nana, [čubar](#), anis, mirođija, lovor) koji olakšavaju varenje. (a, drenjina, <!-- p n = 144 --> grožđa, [dinja](#), kajsija, kuva se još i od šipaka. (S) Bez šećera z od šljiva, jabuka, bresaka, ringlova, [drenjina](#), <!-- p n = 144 --> grožđa, dinja, kajsija, ku z „slatkiji od ukuvane zgusnute šire, od [dudinja](#), grožđa ili drugog voća.“ (S) I za Borda Popović , sitno izrezani komadi tikve i, retko, [dunjja](#). (S) Ovakvom „recelju“ se tokom „pečenja“ dodaje š ene domaće varijante: elda, elsa, elja, [eljda](#), jeda, jejda, jelda. (S) I Hrvati imaju heljdu, a iku kod nas se već dugo koristi i naziv [feferoni](#), a u novije vreme sve više se probija i čili,

Figure 2. Extract of the concordances obtained by the grammar in Fig. 1 from source A6

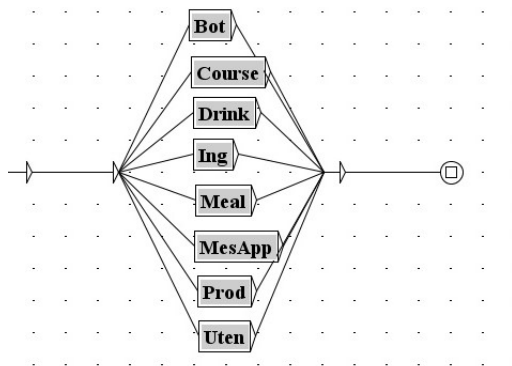


Figure 3. TAG automaton

crnog luka (crni luk, Comp+Bot+...+Food+Alim) i maslinovog ulja (PROD: maslinovo ulje, N, Comp+Conc+Prod+Food). Servirala se ohlađena, s kiselim mlekom (PROD: kiselo mleko, N, Comp+Ek+Conc+Food+Prod+Course).¹⁹

The names of plant ingredients and foodstuffs have been recognised and they have been assigned word class tags (N) and semantic codes. Multi-word lexical units (e.g. *crni luk* [onion]) have been recognised. However, it can be seen from the text that *boranija* [green bean], apart from being the name of a plant, can also be the name of a dish: in Bosnia, *boranija* is the common name for different dishes (here from *spanać* [spinach] and other ingredients). Such examples suggest the need for tagged texts to be additionally analysed so that a certain meaning can be temporally and territorially specified.

We must note that, due to differences in the language of food, the corpus should be broken down into thematic micro-wholes instead of whole texts. Namely, in group A texts, besides documenting the language of cooking, there are significant parts that are “empty” in this respect. On the other

¹⁹ Recipes are also given, and thus we find out that at the time, green bean (*boranija*, Bot+...+Food) was prepared from spinach (*spanać*, Bot+...+Food+DOM=Culinary), rice (*pirinač*, Bot+...+Food), onion (crni luk, Comp+Bot+...+Food+Alim) and olive oil (PROD: maslinovo ulje, N, Comp+Conc+Prod+Food). It was served cold, with soured milk (PROD: kiselo mleko, N, Comp+Ek+Conc+Food+Prod+Course).

hand, the results for a certain search key are scattered throughout a text. Figure 4 shows a histogram of occurrences of the noun *luk* in B3. Almost all the occurrences of this noun are in the first half of the text, and then about 3% towards the end of the text, in the section on preparing vegetable preserves. This indicates that the second half of the sample can be excluded from the search with this key. The breaking up of the text into micro-wholes would lead to a partitioning of the entire text into small sections which would have key words relevant to the section as their meta-data.

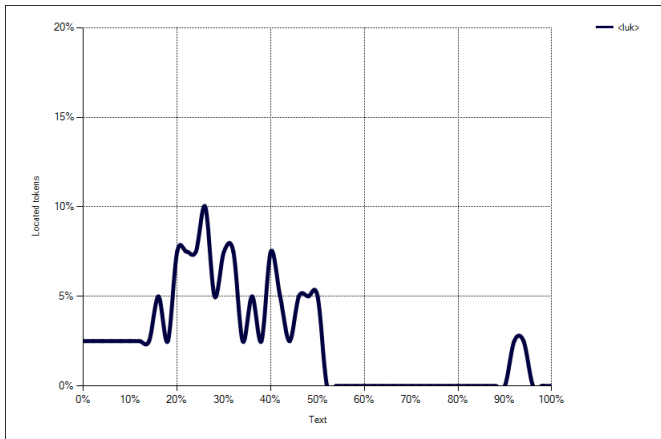


Figure 4. The distribution of occurrences of the noun *luk* in B3

5 Conclusion

The paper presents elements for the design of a corpus of the sub-language of food in Serbian and analyses some of the directions for its development, and also looks at the problems of text selection and annotation. It suggests possible applications which can be developed starting with the language documented in the corpus. In addition to providing searchable material for various disciplines, the construction of such a corpus would provide a basis for compiling an encyclopedia of dietary culture among the Serbian people,

and at the same time a means of transcending various mystifications about culinary traditions.

References

- Banićević, Marta and Magdalena Popović. *Rečnik ugostiteljstva : (srpski-engleski-nemački-francuski-italijanski-ruski) [Hospitality industry dictionary: (Serbian-English-German-French-Italian-Russian)]*. Udruženje naučnih i stručnih prevodilaca Srbije, 2010
- Courtine, Robert. *Dictionnaire de cuisine et de gastronomie*. Larousse, 1986
- Gerhardt, Cornelia, Maximiliane Frobenius and Susanne (eds.) *Ley. Culinary Linguistics*. John Benjamins Publishing, 2013
- Jaćimović, Jelena. “Textometric methods and the TXM platform for corpus analysis and visual presentation”. *Infotheca - Journal for Digital Humanities* Vol. 19, no. 1 (2019): 30–54. https://infoteka.bg.ac.rs/ojs/index.php/Infoteka/article/view/2019.19.1.2_en
- KrsteV, Cvetana and Ranka Stanković. “Old or new, we repair, adjust and alter (texts)”. *Infotheca - Journal for Digital Humanities* Vol. 19, no. 2 (2019), [the same issue]
- KrsteV, Cvetana, Staša Vujičić Stanković and Duško Vitas. “Approximate Measures in the Culinary Domain: Ontology and Lexical Resources”. In *Proceedings of the 9th Language Technologies Conference IS-LT*, 2014, 38–43
- KrsteV, Cvetana, Denis Maurel and Duško Vitas. “Serbian Language Integration in Prolexbase Multilingual Dictionary”. *Infotheca - Journal for Digital Humanities* Vol. 18, no. 2 (2019): 29–52, https://infoteka.bg.ac.rs/ojs/index.php/Infoteka/article/view/2018.18.2.2_en
- Mijo, Kristijan. *Rečnik zaljubljenika u gastronomiju [Dictionnaire amoureux de la gastronomie / Christian Millau]*. Službeni glasnik, 2012
- Mirilov, Ružica. “*Kulinarska terminologija Vojvodine [Culinary Terminology in Vojvodina]*”. PhD. thesis, Filozofski fakultet, Univerzitet u Novom Sadu, 2016
- Montanari, Masimo. *Hrana kao kultura [Food as Culture]*. Sandorf, 2011
- Onfre, Mišel. *Gurmanski um : filozofija ukusa [La raison gourmande / Michel Onfray]*. Gradac, 2002
- Paumier, Sebastian. *Unitex 3.1 User Manual*, Université Paris-Est, 2016. <https://unitexgramlab.org/releases/3.1/man/Unitex-GramLab-3.1-usermanual-en.pdf>

- Radonjić, Danijela. "Kulinarska leksika u savremenom srpskom jeziku [Culinary Lexica in Contemporary Serbian]". PhD. thesis. Filološki fakultet, Univerzitet u Beograd, 2016
- RSJ. *Речник српског језика [Dictionary of Serbian Language]*. Матица српска, 2011
- Slapšak, Svetlana. *Leteći pilav : antropološki eseji o hrani [Flying Pilaf: Anthropological Essays on Food]*. Beograd: Biblioteka XX vek, 2014
- Stošić, Anđela, Marina Milošević, Sandra Spasić, Dragica Dragosav, Ana Đorđević et al. "Working on the Multimedia Document "Al' se nekad dobro jelo"". *Infoteka* Vol. 17, no. 2 (2017): 97–118. <http://infoteka.bg.ac.rs/pdf/Eng/2017-2/infoteka-2017-17-2-5.pdf>, 5
- Vujičić-Stanković, Staša, Cvetana Krstev and Duško Vitas. "Enriching Serbian WordNet and Electronic Dictionaries with Terms from the Culinary Domain". In *Proceedings of the Seventh Global Wordnet Conference*, 2014, 127–132
- Vukov, Lazar. *Imenik jela i pića : na srpskom, francuskom, nemačkom i engleskom jeziku [Directory of dishes and drinks: in Serbian, French, German and English]*. Zadruga, 1954
- Zirojević, Olga. *Istočno-zapadna sofrа: mali kulturno-istorijski i kulinarski leksikon [East-West sofrа: a small cultural-historical and culinary lexicon]*. Beograd: Geopoetika izdavaštvo, 2019
- Балзак, Оноре де. *Голлицаве приче [Les contes drolatiques / Balzac, Honoré de]*. Космос, 1933
- Вивијен, Херберт. "Србија – рај сиромашног човека [Serbia - poor man's paradise]". In *Београд у деветнаестом веку – из дела страних писаца*, 163–199. Библиотека града Београда, 2008
- Витас, Душко and Цветана Крстев, "Оглед из гастрономатике [Experiments in gastronomatics]". In *Теме језикословне у србистици кроз дијахронију и синхронију [Linguistic topics in Serbian through diachrony and synchrony]*, Дражић, Јасмина, Исидора Бјелаковић and Дејан Средојевић, 1–10. Нови Сад: Филозофски факултет, 2016
- Витас, Душко М. "Храна из нежељене поште : (анатомија језика брзе хране [Spam Food: (Anatomy of Fast Food Language)])". In *Српски језик и његови ресурси: теорија, опис и примене*, Ћорић, Божо and Александар Милановић. *Научни састанак слависта у Вукове дане*. Vol. 47, Chapter 2, 21–35. Београд: Међународни славистички центар, Филолошки факултет, Универзитет у Београду, 2018. http://doi.fil.bg.ac.rs/pdf/eb_ser/msc/2018-3/msc-2018-47-3-ch2.pdf
- Вујићић Станковић, Сташа and Весна Пајић. "Формирање доменског корпуса – кулинарска лексика [Preparation of a Domain Corpus – Culini-

- ray Lexical]”. In *Научни састанак слависта у Вукове дане – Српски језик и његови ресурси: теорија, опис и примене*, Vol. 43, Chapter 3, Међународни славистички центар, 2014, 51–59
- Вујићић Станковић, Сташа and Весна Пајић. “Употреба властитих имена у кулинарском домену [The use of proper names in the culinary domain]”. In *Научни састанак слависта у Вукове дане – Српски језик и његови ресурси: теорија, опис и примене*, Vol. 44, Chapter 3, Међународни славистички центар, 2015, 137–142
- Голубовић, Видоје. *Механе и кафане старог Београда [Pubs and taverns of old Belgrade]*. Београд: Службени лист, 2007
- Грђић-Бјелокосић, Лука. “Српска народна јела у Херцеговини и у Босни”. In *Српска народна јела и пића, Књ. 1 [Serbian folk dishes and drinks, book 1]*, Ердељановић, Јован. Београд: Српска краљевска академија, 1908
- Ердељановић, Јован (прир.). *Српска народна јела и пића, Књ. 1 [Serbian folk dishes and drinks, book 1]*. Београд: Српска краљевска академија, 1908
- Игњатовић, Јаков. *Милан Наранџић*. Матица српска, 1949
- Крстев, Цветана and Биљана Лазић. “Глаголи у кухињи и за столом [Verbs in the Kitchen and at the Table]”. In *Научни састанак слависта у Вукове дане – Српски језик и његови ресурси: теорија, опис и примене*, Vol. 44, Chapter 3. Међународни славистички центар, 2015, 117–136
- Марковић, Спасенија-Пата. *Велики народни кувар [Big Folk Cookbook]*. Београд: Народна књига, 1959
- Милорадовић, Софија (прир.). *Обредна пракса - речима о храни : на материјалу из српских говора Војводине [Ritual practice - words about food: on material from the Serbian speeches of Vojvodina]*. Матица српска, Одељење за књижевност и језик, 2014
- Мијатовић, Станоје М. “Српска народна јела (са прилогом о пићима) у Левчу и Темнићу”, In *Српска народна јела и пића, Књ. 1 [Serbian folk dishes and drinks, book 1]*, Ердељановић, Јован. Београд: Српска краљевска академија, 1908
- Огњановић, С. Ф. (прир.). *Ратни кувар у којем су упуства за припремање јела већим делом готовљена без меса а за ове ратне прилике : ручна књига за наше домаћице [A war cookbook in which instructions are mostly for meat-free dishes, and for these war times: a Handbook for our housewives]*. Ујвидек, 1915

- Пелапрат, Анри Пол. *Први кувар света : модерна француска и међународна уметност кувања* [*L'Art culinaire moderne / Henri-Paul Pellaprat*]. Prosveta, 1973
- Петроније, Гај Арбитер. *Сатирикон* [*Satyricon*]. Српска књижевна задруга, 1976
- Поповић Мицина, Катарина. *Илустровани српски "Велики кувар" са 194 слике* [*Illustrated Serbian "Big Cookbook" with 194 illustrations*]. Нови Сад: Српска народна задружна штампариа, 1878
- Портић, Милијанко. *Гастрономски производи* [*Gastronomic products*]. Природно-математички факултет, Департман за географију, туризам и хотелијерство, 2011
- Радуловачки, Љиљана. *Традиционална исхрана Срба у Срему* [*Traditional nutrition of Serbs in Srem*]. Матица српска, 1996
- СТИЈОВИЋ, Рада, Олга Сабо and Ранка Станковић. "Речник САНУ као база термилошких речника (на примеру речника кулинарства) [SASA dictionary as a base of terminology dictionaries (illustrated by culinary lexica)]". In *Међународни научни симпозијум Словенска терминологија данас (Београд, САНУ, 11-13. мај 2016)*, 2016, 109–123
- Тројановић, Сима. *Старинска српска јела и пића* [*Old-Fashioned Serbian Dishes and Drinks*]. Београд: Просвета, 1983
- Фотић, Александар (прир.). *Приватни живот у српским земљама у освит модерног доба* [*Private life in Serbian countries in the dawn of modern times*]. Clio, 2005

6 Appendix

| | |
|----|--------------------------|
| A1 | (Тројановић, 1983) |
| A2 | (Мијатовић, 1908) |
| A3 | (Грђић-Бјелокосић, 1908) |
| A4 | (Радуловачки, 1996) |
| A5 | (Slapšak, 2014) |
| A6 | (Zirojević, 2019) |

Table 5. The group of ethnographic and anthropological texts

| | |
|----|------------------------|
| B1 | (Поповић Мицина, 1878) |
| B2 | (Огњановић, 1915) |
| B3 | (Марковић, 1959) |

Table 6. The group of textbooks