

Anja Nikolić-Hoyt
 Croatian Academy of Sciences and Arts
 Linguistic Research Institute
 Zagreb, Croatia
anhoyt@hotmail.com

Completing an unfinished historical dictionary

1. An introduction to Benešić's *Dictionary of the Croatian Literary Language*

The necessity of the completion of *Rječnik hrvatskoga književnoga jezika od preporoda do Ivana Gorana Kovačića* [The Dictionary of the Croatian Literary Language from the (Croatian) Revival to Ivan Goran Kovačić] (Kovačić being the last out of the 107 Croatian writers whose works are cited in the dictionary), compiled by the renowned Croatian translator, lexicographer, literary critic and writer Julije Benešić, had been kept in mind since the first twelve volumes of the dictionary were published in the late 1980s, some thirty years after lexicographer's death in 1957. The last volume, PROTIVAN-RZATI, was published in 1990. Almost twenty years later, now already fifty years after it was initiated, work on the completion of the dictionary is getting into full swing. Again.

Devised as an anthology of the Croatian Literary Language, Benešić's dictionary was based upon quotations excerpted from works of the 107 most distinguished Croatian authors – chosen by means of somewhat idiosyncratic, i.e. *compiler-specific*, criteria – who were active writers between the mid-nineteenth and the mid-twentieth century. However, this was a rather turbulent period in the history of the Croatian language and literature, as it was marked by significant changes, ranging from the mid-1800s and the establishment of the standard Croatian language to the end of the Second World War and the early beginnings of the Communist era. Owing to the fact that Benešić's dictionary was not normative and exclusive, but rather descriptive and open to the corpus it was based upon, it is also a valuable social and historical monument of the cited period. More clearly, lacking any normative pretensions which would have restrained him from the excerpting of citations from

potentially undesirable, *inappropriate* sources, Benešić put together a genuinely descriptive dictionary that recorded the actual state of affairs, with quotations that appeared just as they had been written down, or published. Aspiring to present a portrait of the Croatian Literary Language in the process of its development, this dictionary

"lists those words which our writers have used in the past one hundred years, words with obsolete contents and forms, words that have become extinct, as well as those which are new. There are no rules or instructions on how writers should write correctly; only examples illustrating usage up until today are given." (Benešić: 1985: XXV).

This implies, among other things, that Benešić's dictionary embodies information about different linguistic and nonlinguistic aspirations in the history of the Croatian language, about its distinctive identity as well as its shared history with the Serbian language. Furthermore, taking into account the fact that Benešić's corpus extends to the middle of the 20th century, i.e. up to the establishment of Socialism in Yugoslavia, it also offers us some insight into the lexical changes that occurred in the Croatian literary language from the beginning of the socialist regime until its fall in the early 1990s. In addition to the above, Benešić's dictionary may also shed light on the lexical changes that occurred during the fascist Independent State of Croatia (1941-1945) as a result of its extremely puristic and rigid language policy. In an attempt to recognize the historical and social relevance of Benešić's dictionary, although by no means radically and comprehensively, one can distinguish between three different socio-historical constellations, each with their own linguistic implications, recorded in the usage of the most distinguished literary writers.

1. 1. Benešić's dictionary on the shared history of Croatian and Serbian

As has already been said, the second half of the nineteenth and the first half of the twentieth century was a very tempestuous and eventful, yet formative and creative period, in which the Croatian standard language was established, and also stabilized. However, compared to the present-day Croatian standard, it was less steady and more open to the lexical impact of the Serbian language, due to shared geographical,

political, legal, ethnographic and other circumstances. As a result, Benešić's dictionary, recording the time referred to in its title, includes words that contemporary Croatian dictionaries often omit, because they are considered to be Serbian forms. But then, Croatian writers of the cited period used words like *artija*, *hartija* 'paper', *ćutati* 'keep silent', *ekser* 'nail', *haos* 'chaos', *parče* 'a small part of something' etc. rather often, and particularly in the late twenties when the new South-Slavic state (Yugoslavia), experienced by many as the fulfillment of a much desired political dream, was established. However, in the course of time, the often vigorous South-Slavic enthusiasm started to abate, while recognition of an exclusively Croatian linguistic identity became prominent only after the breakdown of Yugoslavia and the establishment of the sovereign Republic of Croatia.

Here are some examples of today's Serbian words recorded in Benešić's dictionary as used by renowned Croatian writers:

ćutati, *šutjeti* [keep silent].

... nestade vedrine ... svi ćutali smo (*Bego*). – Tišina je ćutala tako strahovito i sva se okolina tako čudesno ukočila (*Kolar*). – Kestenova krošnje – kraj prašnoga puta. I ravnica plava, što svečano ćuta (*Krleža*). – Huso je ukočen, stegnutih usana ćutao, ćutao kao hrid (*Cesarec*).

ekser, m., *čavao* [nail].

Ona drhtećijem rukama stane sada vaditi ekser po ekser, čiviju po čiviju (*Botić*). – Evo, ovaj (plakat) je pribit samo jednim ekserom, pa ga neće vetar skinuti! – puhao je od ljutine narednik (*Goran*).

párče, n., *malen komad nečega* [a small part of something].

...okosnica posljednje guske sa gdje kojim još parčetom mesa oko kičme (*Jurković*). – Svaki zalogaj, svako parče hljeba, imađase svoju točnu tanku mjeru (*Gjalski*). – Onda je ukrstaro Istrom, tražeći pomoći za otkup onog parčeta djedovine svoje (*Car Emin*). – Eto i on je došao amo po parče osjećaja, po parče života (*Horvat-Kiš*). – A Boga ljubiti je dobro, jer tko njega ljubi i njega se boji, toga i Bog ljubi i daje mu dobro srce i parče sreće (*Krnic*). – ...i dijete bi privinula na grudi, da mu zatim dade komad proje, a kadikad i parče slanine (*Hanžeković*). – Ko parče svrgnute noći o vratu mi zvono visi (*Polić*). – Kažem ti, druže, da sam sav bio njezin, sav do najmanjeg parčeta (*Muradbegović*).

As a result of the strongly divergent tendencies between the formerly integrated languages in the former Yugoslavia, more recent Croatian dictionaries do not record the cited words above. Yet, for some of them they propose more appropriate words, for instance, *Rječnik hrvatskoga jezika* (Šonje, 2000) favors *šutjeti* over *ćutati* (“**ćutati** → šutjeti”), *čavao* over *ekser* (“**ekser** → čavao”), and *komad(ić)* over *parče* (“**parče** → komadić”).

1. 2. Benešić's dictionary on lexical changes in the Independent State of Croatia

Benešić's *Dictionary of the Croatian Literary Language* also lists words which were stigmatized as "newly coined words" that served to differentiate Croatian from Serbian during the fascist Independent State of Croatia (1941-1945). However, Benešić's work disproves such assessments, providing us with information that these words in fact were **not** coined in the Independent State of Croatia, but that they were simply in active use then, having been restored to life from an earlier time. In most cases, the words in question were excerpted from literary works of the late nineteenth and early twentieth century, i.e. long before the Independent State of Croatia had been installed. Here are some examples to illustrate these points:

brzovlak, m., *brzi vlak*, *brzovoz* [express train].

Brzovlak se zadržao na zadnjoj mađarskoj postaji (*Gjalski*).¹

vojarna, *kasarna* [barracks].

Uto dođoše do velike, žuto okrečene zgrade: bila to vojarna (*J. Kozarac*). – Vojarna je pomalo pustjela (*Turić*). – ...zamalo će panduri i špijuni zaređati od tvrđe do tvrđe, od vojarne do vojarne, pa će mnogo srce zakucati življe, jer tko je pred onima nevin (*Lovinac*). – ...mimo golemih vojarnica prođoh (*Horvat-Kiš*).²

Literature:

¹ Because extensive disputes among the reviewers of Benešić's manuscript resulted in the omission of all bibliographical data except the author's name from the dictionary, there is no information about the cited work in the dictionary itself. However, the excerpted material used in the dictionary is stored at the Linguistic Research Institute of the Croatian Academy of Sciences and Arts and available for consultation. Therefore, one can find out that the above quotation is from Gjalski's *Little Stories*, 1894, on p. 152.

² All the quotations date back to the early 20th century, from 1887 to 1919).

Benešić, Julije (1985-1990): *Rječnik hrvatskoga književnoga jezika od preporoda do I. G. Kovačića*, sv. 1-12, Globus i JAZU, I-XII, Zagreb.

Rječnik hrvatskoga jezika (2000), LZ Miroslav Krleža i Školska knjiga, Zagreb.

Samardžija, Marko (2008). *Hrvatski jezik, pravopis i jezična politika u Nezavisnoj Državi Hrvatskoj*, Hrvatska sveučilišna naklada, Zagreb.

zemljovid *zemljopisna karta* [map]

Od novijih učila imadu dva velika zemljovida, scot (računalo) i prirodopisne slike (*Jurković*). – Na stijeni vise dva ogromna i posve požutjela zemljovida: jedan Evrope s Azijom i sa sjevernom Afrikom, a drugi varaždinske županije (*Gjalski*).³

These examples reveal that the proscribed words did not originate in the Independent State of Croatia, but that they were in active use then only as revived words, i.e. lexemes that had been restored to life, for various, mostly non-linguistic reasons. Refraining from a more detailed discussion on language policy in the Independent State of Croatia (1941-1945), when certain words indeed were newly coined (*munjovoz* 'tram', *slikokaz* 'cinema', *voznica* 'ticket') or calqued (*krugoval* 'radio', *cestovna željeznica* 'tram', *dalekopisni stroj* 'telegraph'), it is worth pointing out here that Benešić's dictionary gives better insight into the lexical changes introduced by the Independent State of Croatia, among which newly coined words do not prevail, but rather those which were revived from the literary works of the old Croatian writers (as more recent studies have shown).

1. 3. Benešić's dictionary: before and after the Communist era

Finally, since Benešić's dictionary reflects the pre-Communist era in the history of the Croatian language, it can also assist in reconstructing lexical changes which the Croatian language underwent from the establishment to the break-up of the Communist regime. In other words, resulting from the markedly convergent and unitarian language policy of Communist Yugoslavia, the beginnings of which date back to the late nineteenth century, a significant number of genuinely Croatian words went out of use. But luckily, these words are recorded as distinguishing traits of the Croatian literary tradition to a large extent in quotations found in Benešić's dictionary.

³ The first quotation was taken from Jurković's *Collected stories* (1880, p. 101) and the second is from Gjalski's novel *Under the old roofs*, first published in 1886, p. 42.

However, it is worth mentioning that among these missing words, contrary to expectations that vanished words usually denote missing referents (nouns) or abandoned activities (verbs), there are a fairly large number of abstract nouns as well as highly expressive adjectives and adverbs. Here are just a few examples:

izmicaus and **izmicauz** adv. 'sneaking away unobserved'

Samo da mu se je najest, pa onda izmicaus u bekrijanje (*Kosor*). – Da si vidio, što je «izmicauz», kad švabo dade pleća! (*Cepelić*).

jutroske, adv.; *jutros*. 'this morning'

Jutroske je rano bez kraja i konca brbljo vrutak jureć niza stijenje vlažno (*Nazor*).

mimogrece, adv.; *mimogred, usput, uzgred*. 'by the way'

Zatim od Grka ... prijeđe na srednji vijek, gdje mimogrece spomenu i fiziokratičku školu (*Kovačić*). – I vrzla mu se glavom slična nekakva misao – dakako tek letimično i mimogrece (*Gjalski*).

neboderan 'sky-scraping'

Na visoku proplanku, što se stoljetnom šumom uznosita hrašća, nebodernih jela i omorika veže o plavetno timorje veličanstvenoga Vlašića (*Hanžeković*).

nebogreban 'sky-scratching'

Golema zgrada ... sada je sva još zagrađena nebogrebnim lazilima (*Horvat-Kiš*).

nebotičan 'sky-touching'.

Nebotični hrastovi stoje ponosito jedan do drugoga kao neka gorostasna vojska (*Tombor*). –... ako i nađu nebotičnu jelu, ne mogu je izvući iz šume zbog duljine, dok je ne prikrate (*Goran*).

Taking into account various issues discussed briefly in the previous paragraph, it is no exaggeration to say that Benešić's dictionary is a document about the stability of the Croatian lexis, as well as its variability resulting from significant socio-historical changes. The social and historical relevance per se of the dictionary is beyond any doubt striking. Needless to say, much credit goes to the corpus, which is both fascinating and intriguing.

2. Problems deriving from the completion of an unfinished historical dictionary

Two years ago, when the actual project concerned with the completion of a historical dictionary of the Croatian literary language started, I was faced with seventy or so dusty boxes filled with thousands of barely readable cards, as the attached example confirms.

1854.

✓ stîžnik a mîder

u stîmi'je naprîao h' u o'ci, e'zi
 u'vita stîpuka, u'ci se pi' h'ajitak
 — nekako stîpuk kraj — stîpuk
 prîvija oko krasnog stîpuk j'oj
 struka

1886. Gjalma. Pr. st. H. 338

Despite my despair verging on anger with the ungainly scribes I began to work with enthusiasm and patience, starting with the endless letter S. After mastering the necessary paleographic skills, however, I came across other problems, some of which I'll try to discuss within this paper. All in all, two years later or so, in spite of the fact that it is still a "work in progress" and apparently will remain "in progress" for years to come, I must admit that working on this project has already proven to be worth the effort. However, *the great labor is yet to come!*

What is more, at the time of its inception, Benešić's dictionary was conceived as a dictionary of the contemporary Croatian literary language. It goes without saying that what was then considered contemporary lexis is quite different from that of today, and this fact disrupts today's reading as well as writing of Benešić's distinguished dictionary.

On the other hand, Benešić's dictionary was never compiled according to a clear set of (lexicographic) rules, so the present-day project is confronted with

numerous problems, beginning with the conciliation of two different aspirations: the goal of finishing the dictionary the way the author would have done it, and the urge to amend his misjudgments.

First of all, taking into account the lack of space, Benešić and his reviewers for the most part thought that quotations followed by the name of the author would suffice without including other bibliographic information, such as the title of the quoted work, the edition, and the page number. Although quoting these additional, albeit necessary data, will undoubtedly lead to the revision of the already published volumes, it would increase the potential application of Benešić's dictionary to the utmost. However, the optimal solution, which would also accelerate the work and reduce the number of copying mistakes, should include the digitalization of all the literary sources, i.e. their transformation into a retrievable corpus. For the time being, however, we are stuck with difficult handwriting, a plethora of transcription errors, and the fact that only about 40 percent of the sources are available for consultation. Specifically, a large majority of the sources are rare books, mostly first or early editions that are not always accessible.

In addition to the already mentioned problem of insufficiency of data concerning the sources from which quotations have been excerpted - it is not sufficient for the user to know that a certain quotation derives from the nineteenth-century author August Šenoa, for instance, for there are about twenty or so of Šenoa's works cited - a major objection to Benešić's dictionary pertains to its inconsistency, whether dealing with meaning or grammar. The arguments made by Benešić in defense of his view of what should be defined - for example, that he does not define words which are common or transparent (Benešić, X) - have no merit. Furthermore, when defining transparent word formations, he uses inconsistent metalanguage, as is represented by the following adjectives: *bezbojan* 'colorless' - *što nema boje* 'what has no color'; *bezlistan* 'leafless' - *koji nema lista* 'which has no leaves'; *besputan* 'roadless, pathless' - *gdje nema puta* 'where there is no road; path', *bezruk* 'armless' - *bez ruke* 'without arms'; *besuzan* 'tearless' *što nema suza* 'what has no tears'; etc. As can be seen from these examples, Benešić defines the same type of adjectives, descriptive adjectives, using various defining styles. As far as the overall grammatical coding is concerned, Benešić's dictionary provides very basic grammatical information, often with a lack of consistency. However, given the specific nature of such a dictionary, grammatical data are of marginal relevance in any case.

Perhaps most intriguing and interesting, but also difficult to deal with, are idiomatic usages pertaining to a single author. Even though idioms are considered to be fixed expressions which have a meaning that is different from the meanings of its composite elements, a large number of idiomatic expressions found in Benešić's dictionary are not truly fixed. On the contrary, they are highly variable in reflecting a single author's experience of reality, and they do not usually occur outside his or her corpus. However, within such an author-specific corpus, they represent in fact a certain detachment from the more common idiomatic forms. For example, *staviti na tortu* 'to put on a cake', from Šenoa's novel *Vladimir* (*Majka bi ga stavila na tortu; ako razbije nekomu kamenom prozor, tvrdi majka, da joj je sin genije*. [His mother would put him on a cake; if he threw a rock through someone's window, she would say that he was a genius.]), is an author-dependent variation of a familiar idiomatic expression *staviti koga na pijedestal* 'to put/set someone on a pedestal'.

Last of all, the most strongly stressed principle, and at the same time the most consistently conducted principle, in Benešić's dictionary is the chronological ordering of citations, from the oldest to the most recent. Still, it is not always easy to meet the requirements due to the fact that there is a significant disproportion between data written on the cards and bibliographical items used as sources as listed at the beginning of the first volume. In comparing these two, one notices that the difference between editions sometimes exceeds two or three decades. For example, according to the List of excerpted authors and their works, *Stories*, by Janko Leskovar, was published in 1917. However, all the citations recorded on the handwritten cards were based upon the newer 1944 edition. Unfortunately, the two editions differ in terms of both orthography and lexis.

3. Benešić's dictionary in both historical and contemporary lexicographical paradigms

Finally, Benešić's dictionary contains elements that place it in both historical and contemporary lexicographical paradigms. As in traditional historical dictionaries, each word is described via chronologically ordered quotations of actual usage, so that the reader can follow gradual changes of form and meaning in the language over the course of one hundred years. Although Benešić's dictionary was never truly systematic and thorough, as all distinguished dictionaries should be, but in terms of procedure, two centuries after Samuel Johnson, Benešić was in fact using the same

method. Having selected writers whom he thought were the most outstanding literary authorities, he went through their books and marked everything that he thought could be used. His excerptors (three graduates of Croatian language and literature) copied the words, including the surrounding passages, onto blank cards, following the directions of the editor and their own intuition. The cards were then sorted alphabetically by headword, and Benešić decided, according to the illustrativeness of the ‘material and his personal preference’, which ones he would use. After defining the corpus, he had to decide how many senses to recognize for each word, write the definitions, and provide some grammatical information. However, unlike Johnson, Benešić aimed to produce a descriptive dictionary; he referred to it as *informative*, not normative, as Johnson’s prescriptive dictionary was, due to the contemporary concerns about the state of the language.

On the other hand, Benešić’s cards were very much like James Murray’s slips of paper, with the word in the top left-hand corner and full bibliographical details. However, Benešić’s entries were far more simple and less revised than the polystructured entries in the *New English Dictionary on Historical Principles*, which showed the ‘form-history’ of each word, as well as a thoroughly documented history of its meanings.

Finally, in light of contemporary concerns about lexicographical evidence, every decision in the dictionary is, so to say, corpus-based - of course, only if we regard a corpus as any body of text collected with the aim of analyzing it. Consequently, although Benešić did not use a language corpus as a basis for his dictionary, all the senses that he recognized were derived from his collection of evidence, i.e. quotations (in the form of paper slips). And this means that, when a certain word appeared in quotations only in their second and/or figurative meaning(s), Benešić would record no more than those meanings. In other words, the fact that the conventional meaning is not recorded indicates that Benešić had no evidence of it.

However, neither do all the quotations provide sufficient evidence of the meanings of words nor do they always illustrate the changes and developments in meaning in a reasonable manner. Here, the difference between the corpus evidence and the quotation slips is most striking. Let me give some examples that show the *escapability* of the information needed to form a definition from the collected quotations:

strankinja, f. 'a foreign woman'

Ona je strankinja. 'She is a foreigner.' (Šop)

strast, f., *veoma snažna goruća želja; požuda*. 'passion'

Strast je majka svega što je lijepo. U njoj se rodiše sva najuzvišenija djela. 'Passion is the mother of everything beautiful. All the most exalted works were born of it.' (Andrijašević) – *Lice je mirno, na njem nema strasti.* 'The face is peaceful, there is no passion in it.' (Šenoa). – *Lov je moja strast.* 'Hunting is my passion.' (Tomić).

On the one hand, the corpus gloss may not not have been sufficiently explanatory. On the other hand, Benešić's quotations were taken from literary language, which is sometimes, owing to the fact that literature is an art of words, evoked like other arts by momentary inspiration, which is difficult to reconstruct. Whatever its shortcomings, Benešić's dictionary provides at least some insights into the use of the Croatian language in the most formative period of its history and is, therefore, a valuable tool, as well as a distinct achievement in Croatian lexicography.

Literature:

- Benešić, Julije (1985-1990): *Rječnik hrvatskoga književnoga jezika od preporoda do I. G. Kovačića*, sv. 1-12, Globus i JAZU, I-XII, Zagreb.
- Rječnik hrvatskoga jezika* (2000), LZ Miroslav Krleža i Školska knjiga, Zagreb.
- Samardžija, Marko (2008). *Hrvatski jezik, pravopis i jezična politika u Nezavisnoj Državi Hrvatskoj*, Hrvatska sveučilišna naklada, Zagreb.