

The Terms We Share or Don't Want to Share – On Multiple Entries in *STRUNA*

Bruno Nahod, Perina Vukša Nahod

Institute of Croatian Language and Linguistics, Zagreb, Croatia

ABSTRACT

Through the decades of working on processing Croatian terminology in STRUNA we have identified two problems that can generate an extremely confusing situation for the end-user. On the one hand, we have homonyms, where the same term is used to indicate multiple concepts. This problem was anticipated since it is well-known that certain terms can have different or slightly different meanings in various fields of knowledge or even in the same field. We handle the homonyms by allowing multiple entries in STRUNA. On the other hand, fairly often, we come upon the case where a new entry was introduced into STRUNA even though the term with a semantically identical definition already existed. In anticipation of this problem, we have, from the very beginning of STRUNA, implemented the option of accepting the existing terminological entry and associating it with the ongoing project. Unfortunately, for various reasons, editors rarely use this feature. Therefore, this issue will have to be addressed in the future. In this paper, we will present these problems experienced by the end-user of STRUNA and suggest possible solutions implemented in a certain project (field) and/or interfiled solutions.

Key words: terminology management, Croatian terminology, end-user satisfaction

Introduction

Continuous joined efforts of terminologists, field experts, and language experts in the national terminological database *STRUNA* since 2009 has yielded a term bank comprising 31 domains produced through 31 individual projects managed by field experts, partially financed by the Croatian Science Foundation and partly by various research institutions in Croatia. Even though both *STRUNA* as a Croatian term base and the good practices of terminology processing and management are well known^{1–3}, various problems and challenges in terminology management regularly emerge.

In this paper, we will present two reoccurring problems in processing terminological units that we have identified as possible causes of misunderstanding or misleading the end-user. The end-user's satisfaction is not a novelty and has been recognized as one of the fundamental aspects of managing information systems which terminological databases are a big part of^{4,5}.

One of them is the problem of homonymy, namely the cases where the same term designates more than one concept within the same domain or, more frequently, within

different domains. The other one is the case where multiple entries were produced by the experts of the different domains where both the term and the definition of the concept are (basically) the same.

Even though the traditional terminological theory⁶ assumes the one-term-one-concept principle, praxis has shown us that the rule is often untenable. Deviation from the principle is justified if sufficient arguments exist for double or multiple entries. In those cases, it is not a question of polysemy but of homography⁷. On the other hand, two entries in different domains with the same term and minimal differences in definition can be considered unjustified, which implies that one definition is optimal while the others are suboptimal. In the long term, this is one of the problems that will regularly appear in a large database because we can assume that new entries will be processed as part of sub-domains of the already processed domains.

We will show how end-users are currently experiencing such examples when searching *STRUNA* and how some experts solved these problems while working on their terminology.

Current situation in STRUNA

Often, when a *STRUNA* end-user searches for a particular term, he can get several results, which can be confusing because it is hard to know whether the multiple matches of the query are the result of ambiguity within one or more domains or a case of multiple entries of the same terminological unit (Figure 1).

A good model of terminology processing includes several stages. First of all, it is important that during the initial workshops, domain experts become familiar with the workflow and principles that are being followed in *STRUNA*, which should be implemented at all levels in all domains. They can choose between creating an abecedary or a conceptual system depending on their original starting strategy. The next stage is extracting existing terms and definitions from available sources and adapting them according to acquired terminological principles. The next step is to choose the optimal terms as well as corresponding linguistic determinants, definitions, equivalents in foreign languages, synonyms in Croatian, subordinate terms, abbreviations in Croatian or a foreign language, symbols, equations, formulas, links, attachments, notes, contexts, and sources of terms, definitions, or contexts. This is followed by terminological processing carried out by a terminologist and proofreading carried out by language experts.

It is recommended that several domain experts from different subfields participate in the project and an editor-in-chief who, together with a terminologist and a (Croatian) language consultant, process and refines the terms and definitions in their designated phases of terminology processing.

One of the previously mentioned problems that may arise during this stage is twofold: the term that the editor wants to enter in *STRUNA* has already been processed and entered as part of another project and is publicly available in *STRUNA* or has already been entered as part of another project. The editor must work closely with the project editor-in-chief and terminologist at this stage so that it can be determined if additional entries are necessary.

Various tools to avoid unnecessary multiple entries are implemented into *STRUNA*'s CMS (Content Management System). The system automatically flags the term when it already exists in *STRUNA*. The editor gets a notification on his screen that a duplicate is about to be added and that further action is needed before continuing.

If it is determined that an identical terminological unit has already been processed and published in *STRUNA* as part of another project and that the term, synonyms, definition, and possible notes and contexts correspond to the domain that is currently being processed, the simplest solution is to “like” the term. The colloquial term “like” represents a process of adopting a terminological unit into an existing project, indicating that the editor, editor-in-chief, and terminologist are satisfied with the entered data. This feature is also presented in the front end of , where both the original project that entered the terminological unit and all the other projects that liked it are listed as “owners”. An example of the result of adopting a terminological unit can be seen in Figure 2 where the sound (*zvuk*) that physicists originally entered was adopted by linguists.

If the term signifies an entirely different concept or that the editor of a domain is not satisfied with the existing terminological unit, for example, due to inaccuracy or the scope of the definition, the absence of subordinate concepts, or any additional information that is crucial for the ongoing project, that terminological unit is re-entered into *STRUNA* and processed in accordance with the meaning in the domain that is currently being processed. The result of this, from the end-user's standpoint, is a duplicate in the search results.

If the editor wants to enter a term that has already been entered in his or her project, it should be determined whether it is the same entry that can be associated with two or more subdomains, or if it is a case where the same term is used to represent a different concept in different subdomains.

In the first case, the best solution is one entry and, consequently, one definition, which has been edited by the

Rezultati pretraživanja za: OKVIR 1 2 3 » »»

okvir element strukture trupa zrakoplova koji mu daje oblik	zrakoplovstvo, raketna i svemirska tehnika
okvir osnovni nepokretni dio koji nosi vratno ili prozorsko krilo i pričvršćuje se u zidni otvor	drvena tehnologija
okvir jedna od mirnih slika koje tvore pokretnu sliku	sigurnosne i obrambene znanosti
okvir temeljni dio vatrenoga oružja na koji se u proizvodnji postavljaju ostali dijelovi	sigurnosne i obrambene znanosti

Fig. 1. Search results for the term okvir (frame).

zvuk <small>im. m. jd.</small>	
definicija	val frekvencije od približno 20 Hz do 20 kHz na koji je osjetljivo ljudsko uho, nastaje zbog periodičnoga titranja čestica u elastičnome sredstvu
istovrijednice	engleski: sound
razredba	polje: fizika grana: opća i klasična fizika projekt: Izgradnja, odabir i usuglašavanje hrvatskoga nazivlja u fizici, Hrvatsko jezikoslovno nazivlje (JENA)

Fig. 2. Example of how the end-user sees when the terminological unit is “liked”.

joint effort of two or more editors, the editor-in-chief and terminologists, and where the optimal edit is presented, one which encompasses all the different nuances for the terminological unit to be suitable for each subdomain of interest. Additionally, something that can significantly help in finding an optimal solution is the field “notes,” in which it is possible to explain what is important in a specific area or what may differentiate them.

If, for well-argued reasons, this kind of solution is not possible, for example, where a completely different definition is given for two or more subdomains with the same term or within different theories or schools, the traditional terminological principles must be deviated from, i.e., each term is edited with a different definition, and this results in two entries (Figure 3).

Such problems appeared almost with each project in *STRUNA*, allowing us to use examples from several projects to show our proposed solution.

Croatian Linguistic Terminology - Jena

Croatian Linguistic Terminology - Jena⁸ is one of the more recent projects of *STRUNA*, and it was started due

to the need to standardize Croatian linguistic terminology. It lasted from May 2019 to December 2020 as part of HRZZ financing. Following the project's official end date, Jena was implemented as an internal project of the Institute for Croatian Language and Linguistics. The main reason for a follow-up project was the relatively short timeline of the original project and the aspiration of the main editor-in-chief as well as all of the collaborators, to fill in all the gaps in the conceptual system (which solely existed because of a lack of time) and to keep working on linguistic terminology to keep everything up to date. This unique praxis is a good model of sustainability. It would have been ideal if all previous projects processed in *STRUNA* had implemented something similar to ensure the refinement of already entered terminological units and especially the entry of new ones. As part of the project, 2831 terminological units were processed. A large part of the linguistic terms can be associated with different subfields, and several results of such overlaps are visible both within the project itself and with other projects. In addition, there are terms already entered in *STRUNA* that have entirely different meanings in linguistics. The following is an example of how the above examples are processed within Jena.

Rezultati pretraživanja za: POLJE	
polje prostor u kojemu djeluje koja sila	strojarstvo
polje vektorska veličina pridružena svakoj točki prostorvremena	fizika
polje tijelo u algebarskome smislu u kojemu je operacija množenja komutativna	matematika
polje spiralno ispuščenje u unutrašnjosti cijevi vatrenoga oružja	sigurnosne i obrambene znanosti
polje dio videosignala koji sadržava naizmjenične vodoravne crte	sigurnosne i obrambene znanosti

Fig. 3. Example of a multiple entries of the same term in one project – polje (field).

Terms edited in the projects preceding Jena

By reviewing and exporting the terms from previous projects (domains) whose terminology has been processed within *STRUNA* so far (e.g., anthropology, librarianship, defectology, etc.), it was observed that some terms could be part of the linguistic terminology. In cases where the definition and all components were acceptable to the linguistics team, the terminological units were added to the linguistics project. Thus, for example, the terms language games, language policy, language typology, basic dictionary, and endangered language, which were processed during the project covering anthropological terminology⁹, are adopted to the Croatian linguistic terminology in the same form, which is visible to the user when searching for that term in (Table 1).

New entries due to a different definition

Editors in Jena decided to enter a new term when the term in *STRUNA* has an entirely different definition in their domain. In these cases, since terms designate completely different concepts, homonymy (from the end-user’s point of view) is unavoidable.

For example, the term equivalence exists in *STRUNA* because it was processed as part of the Cartographic-Geoinformatics Dictionary project. Still, it does not match the meaning within traductology and contact linguistics, and the processor processes the term within his subfield, which leads to homonymy (Table 2).

Often, we are faced with multiple entries of the same term, which appear as several homonyms to an end-user. For example, the editor from the subfield of cognitive linguistics entered the term *izvor* (source), which already has two entries in *STRUNA*. This term is included in geodesy (Cartographic-Geoinformatics Dictionary) and civil engineering (Systematization of Croatian Civil Engineering Terminology), and in all cases, it is rightly defined differently (Table 3).

New entries due to a disagreement

Jena editors were observed entering a new term when they were not fully satisfied with the existing definition or other processed properties of the already edited terminological unit in *STRUNA* because they lacked essential information. For example, the editor of the etymological subdomain in linguistics considers that the definition of the term *prajezik* (protolanguage) within anthropology is acceptable. Still, for etymology, it is essential to highlight genetic kinship (Table 4). In cases like this, we are dealing with the “same” concept that is edited in two different ways¹⁰. In addition to expanding the definition, additional information that did not exist in the previous edit was added, for example, a deprecated term and the equivalent in Russian.

Similarly, the editor of pragmlinguistic terminology considered that the term *teorija govornih činova* (speech act theory), which was previously processed in anthropology, should be defined to fit into the pragmlinguistic

TABLE 1
PROCESSING OF THE TERM ‘LANGUAGE POLICY’ IN *STRUNA*

jezična politika	
definition:	<i>djelatnosti kojima institucije nastoje nadzirati i mijenjati jezičnu praksu ili ideologiju</i> (activities by which institutions try to monitor and change language practice or ideology)
English:	language policy
note:	<i>Ove djelatnosti u prvome su redu usmjerene na razvoj, očuvanje, uporabu, učenje i širenje nacionalnih idioma te potvrđivanje njihova službenoga statusa. U širem smislu odnose se i na druge jezične resurse u državi (nestandardne idiome, idiome nacionalnih manjina ili useljenika), kao i na učenje i širenje drugih nacionalnih idioma (tzv. učenje stranih jezika).</i>
field:	<i>etnologija i antropologija</i>
branch:	<i>antropologija</i>
project:	<i>Izgradnja temeljnog nazivlja u antropologiji, Hrvatsko jezikoslovno nazivlje</i>

TABLE 2
COMPARISON OF PROPOSED EDITS OF THE TERMS ‘EQUIVALENCE’ IN *STRUNA* IN GEODESY AND LINGUISTICS

term	definition	equivalents	domain
<i>ekvivalentnost</i>	<i>vjernost prikaza površina na karti umanjenih u skladu s mjerilom karte</i> (the faithfulness of the areal relationships on the map reduced in accordance with the scale of the map)	English: equivalence German: <i>Gleichwertigkeit</i> French: <i>équivalence</i>	geodesy
<i>ekvivalentnost</i>	<i>svojstvo prijevoda da su jezične, paradigmatске, stilske i tekstne značajke polaznoga teksta očuvane u ciljnome tekstu</i> (the property of the translation that the linguistic, paradigmatic, stylistic and textual features of the source text are preserved in the target text)	English: equivalence German: <i>Äquivalenz</i> French: <i>équivalence</i> Russian: <i>эквивалентность</i> Sweden: <i>ekvivalens</i>	linguistics

TABLE 3COMPARISON OF PROPOSED EDITS OF THE TERM 'SOURCE' IN *STRUNA* IN GEODESY, CIVIL ENGINEERING, AND LINGUISTICS

term	allowed term	definition	synonyms	note	domain
<i>izvor</i>	–	<i>organizacija od koje su podatci preuzeti ili način na koji su podatci dobiveni</i> (the organization from which the data was obtained or the manner in which the data was obtained)	English: source German: <i>Quelle</i>	<i>U analizama pogrešaka izvor se najčešće odnosi na izvore pogrešaka i nesigurnosti.</i>	geodesy
<i>izvor</i>	<i>urelo</i>	<i>mjesto na kojemu voda izbija na površinu zemlje</i> (the place where water breaks out on the surface of the earth)	English: spring German: <i>Quelle</i>	–	civil engineering
<i>izvor</i>	–	<i>semantička uloga koja označuje početnu točku nekoga lanca radnje</i> (a semantic role that marks the starting point of a chain of actions)	English: source German: <i>Quelle</i> French: <i>source</i> Russian: источник	<i>Primjerice, u rečenici Ivan je Marku dao poklon. Ivan je izvor, ali ujedno i agens.</i>	linguistics

TABLE 4COMPARISON OF PROPOSED EDITS OF THE TERMS 'PROTO-LANGUAGE' IN *STRUNA*

prajezik <i>im. m. jd.</i> (n. sg. masc.) definition: <i>potvrđeni ili pretpostavljeni predak jednoga poznatog jezika ili više njih</i> (a confirmed or assumed ancestor of one or more known languages) allowed term: <i>jezik predak</i> English: proto-language German: <i>Ursprache</i>	prajezik <i>im. m. jd.</i> (n. sg. masc.) definition: <i>jezik iz kojega su se razvili genetski srodni jezici</i> (language from which genetically related languages developed) allowed term: <i>jezik predak</i> deprecated term: <i>jezik-predak</i> English: proto-language German: <i>Ursprache</i> Russian: <i>праязык</i> field: <i>filologija</i> branch: <i>poredbeno i historijsko jezikoslovlje</i> project: <i>Hrvatsko jezikoslovno nazivlje</i>
---	--

context; therefore, a new entry of the "same" term was made¹¹. (Table 5)

Among such examples there are many that are processed within the domain of anthropology, i.e., language myth, creole language, natural language, and artificial language, which were also processed by the editor of the subfield of sociolinguistics, or linguistic universality, semiotics, semiology, and stereotype, which were processed by the editor of the subfield of pragmalinguistics. Understandably, editors want to define terms within their subfield and harmonize them according to their knowledge and understanding of a specific concept. On the other hand, especially while keeping the end-user in mind, it is necessary to harmonize on the level of the whole term base to avoid unnecessary multiplication. The prime example of such a case is the multiple entries of the term 'language myth', where there are not only two definitions practically the same (Table 6), but they are using the same reference as a source of the definition.

Terminological units entered in Jena

Acceptance and/or re-edits

During the terminological processing within the project Jena, when entering new units from their subfields,

many editors noticed that editors from other linguistic subfields had already entered some terms. They would undoubtedly have been defined differently on their own. Still, in cooperation with the editor, an attempt was made to enter only one term and one definition. A common solution was reached that includes all the details important for individual subfields. This is the optimal workflow according to the terminological guidelines and principles we try to follow in *STRUNA*.

For example, the terms first palatalization, second palatalization, third palatalization, liquid metathesis, weak semivowel, and strong semivowel share the subfields of language history, dialectology, and etymology, where the editor either accepted the existing definition or refined the existing one in collaboration with a colleague.

New entries due a different definition

Editors enter a new term when it is semantically different from the term already entered in another subfield¹², for example, the term *imenovanje* (naming) in the subfields of onomastics and pragmalinguistics (Table 7) or the *domena* (domain) in the subfields of terminology and cognitive linguistics (Table 8), where both of the pairs were entered as unique terminological units.

TABLE 5

COMPARISON OF PROPOSED EDITS OF THE TERMS 'SPEECH ACT THEORY' IN *STRUNA*

<i>teorija govornih činova</i>	<i>teorija govornih činova</i>
definition: <i>teorija o uporabi jezika kao djelovanju</i> (theory of language as action)	definition: <i>proučavanje govornih činova utemeljeno na opisivanju mogućnosti čovjekova djelovanja s pomoću jezičnih obrazaca</i> (the study of speech acts based on describing the possibilities of human action with the help of language patterns)
English: speech act theory	English: speech-act theory
subordinated concept: <i>performativ</i>	German: <i>Theorie der Sprechakte</i>
	French: <i>théorie des actes de langage; théorie de l'acte de parole</i>
	Russian: <i>теория речевых актов</i>
	Sweden: <i>talhandlingsteori</i>
note: <i>Teorija govornih činova proizlazi iz rada Johna L. Austina i Johna Searlea koji su proučavali kako govornici upotrebljavaju jezik da bi postigli željeni učinak te kako slušatelji pronađu u značenje izrečenoga. Osnovna je pretpostavka da svakim izričajem govornici izvode jedan društveni čin ili više njih. Iako je začeto u okrilju filozofije jezika, proučavanje govornih činova danas čini potpodručje međukulturne pragmatike. Teorija je doživjela kritiku u lingvističkoj antropologiji zbog univerzalističkih tendencija i zanemarivanja društvenoga konteksta.</i>	note: <i>Utemeljitelj teorije govornih činova britanski je filozof John L. Austin 60-ih godina 20. stoljeća. Postavke je teorije iznio u knjizi Kako djelovati riječima (How to do things with words), a proširio ju je američki filozof John R. Searle. U osnovi je teorije govornih činova tumačenje da jezik nije samo sredstvo iznošenja informacija i opisivanja nego i način čovjekova djelovanja u društvu i stvaranja promjena.</i>
field: <i>etnologija i antropologija</i>	field: <i>filologija</i>
branch: <i>antropologija</i>	branch: <i>opće jezikoslovlje (lingvistika)</i>
project: <i>Izgradnja temeljnog nazivlja u antropologiji</i>	project: <i>Hrvatsko jezikoslovno nazivlje</i>

TABLE 6

COMPARISON OF PROPOSED EDITS OF THE TERMS 'LANGUAGE MYTH' IN *STRUNA*

<i>jezični mit</i>	<i>jezični mit</i>
definition: <i>rasprostranjeno netočno mišljenje o nekome idiomu ili jeziku općenito</i> (a widespread incorrect opinion about an idiom or language in general)	definition: <i>široko prihvaćeno, ali netočno mišljenje o određenome jeziku ili o jeziku općenito</i> (widely accepted but incorrect opinion about specific language or about language in general)
source reference: Trask (2005.)	source reference: Trask (2005.)
English: language myth	English: language myth
	German: <i>Sprachmythos</i>
	French: <i>mythe de la langue; mythe linguistique</i>
	Russian: <i>языковой миф</i>
	Sweden: <i>språkmyt; språklig myt</i>
field: <i>etnologija i antropologija</i>	field: <i>filologija</i>
branch: <i>antropologija</i>	branch: <i>opće jezikoslovlje (lingvistika)</i>
project: <i>Izgradnja temeljnog nazivlja u antropologiji</i>	project: <i>Hrvatsko jezikoslovno nazivlje</i>

These are the cases that can be the most confusing for the end-user because, in the *STRUNA* search engine, only the field and branch of the project's domain are visible; on the other hand, the subfield, which is self-determined by the project editor for each project, is not visible.

The coining of the new term

One of the possible solutions to avoiding apparent homonymy is the coining of new terms. When possible, the editors created new terms in the linguistic terminology. In etymology, for example, the terms *konsonantizam* (consonantism) and *vokalizam* (vocalism) are used for "consonant" and "vowel system of a given word", while in dialectology these terms are used for the "consonant" and

"vowel system of a given speech" (Table 9). To avoid homonymy, etymological terms were entered as *konsonantizam riječi* (word consonantism) and *vokalizam riječi* (word vocalism)¹³.

Unavoidable ambiguity within a subfield

Sometimes it is impossible to avoid ambiguity within the same field because it is a question of two different concepts designated by the same term. For example, *etimologija* (etymology) is entered twice within the subfield of etymology, and *taksonomija* (taxonomy) within the subfield of terminology (Tables 10 and 11). The entry of such examples is necessary and justified.

TABLE 7
COMPARISON OF PROPOSED EDITS OF THE TERM 'NAMING' IN *STRUNA*

<p>imenovanje <i>im. sr. r. (n. ntr.)</i></p> <p>definition: <i>postupak kojim se imenuje osoba, mjesto ili predmet</i> (the procedure by which a person, place or object is named)</p> <p>source reference: <i>Osnoven sistem i terminologija na slovenskata onomastika</i> (1983)</p> <p>deprecated term: <i>čin imenovanja; proces imenovanja; denominacija</i></p> <p>English: naming; name giving</p> <p>German: <i>Namensgebung</i></p> <p>French: <i>nomination</i></p> <p>Russian: <i>акт номинации</i></p> <p>Sweden: <i>namngivning</i></p> <p>note: <i>Jednorječni naziv imenovanje obuhvaća cjelokupan postupak imenovanja (uključujući i sam čin), te je stoga prikladniji za uporabu od dvorječnih naziva čin i postupak imenovanja. Naziv se denominacija nije ustalio u hrvatskoj onomastici.</i></p> <p>field: <i>filologija</i></p> <p>branch: <i>opće jezikoslovlje (lingvistika)</i></p> <p>project: <i>Hrvatsko jezikoslovno nazivlje</i></p> <p>subfield: <i>onomastika</i></p>	<p>imenovanje <i>im. sr. r. (n. ntr.)</i></p> <p>definition: <i>egzercitiv kojim se komu ili čemu dodjeljuje kakvo ime ili se koga postavlja na kakvu dužnost</i> (An exercitive by which someone or something is assigned a name or someone is appointed to a position)</p> <p>source reference: Austin (2014)</p> <p>English: naming</p> <p>German: <i>Ernennung</i></p> <p>French: <i>nomination</i></p> <p>Russian: <i>именование</i></p> <p>note: <i>U Austinovoj teoriji govornih činova imenovanje (u užemu smislu ceremonija dodjeljivanja imena komu pri crkvenome krštenju ili pri krštenju broda i sl., a u širem smislu proglašavanje koga po imenu kao nositeljem kakve dužnosti) jedan je od temeljnih oblika govornih činova koji, da bi bio uspješan, mora zadovoljiti uvjete prinjerenosti.</i></p> <p>field: <i>filologija</i></p> <p>branch: <i>opće jezikoslovlje (lingvistika)</i></p> <p>project: <i>Hrvatsko jezikoslovno nazivlje</i></p> <p>subfield: <i>pragmalingvistika</i></p>
---	---

TABLE 8
COMPARISON OF PROPOSED EDITS OF THE TERMS 'DOMAIN' IN *STRUNA*

<p>domena <i>im. ž. r. (n. fem.)</i></p> <p>definition: <i>područje znanja, disciplina, proces proizvodnje ili metoda u kojima se određeni pojam upotrebljava</i> (a field of knowledge, discipline, production process or method in which a certain term is used)</p> <p>source reference: Bessé <i>i sur.</i> (1997)</p> <p>English: domain</p> <p>German: <i>Domäne</i></p> <p>French: <i>domaine</i></p> <p>Russian: <i>область; сфера; домен</i></p> <p>Sweden: <i>domän</i></p> <p>note: <i>Sve domene utvrđuju i po potrebi mijenjaju stručnjaci. Domene imaju jednostruk ili višestruk ustroj koji se uglavnom odražava mrežom definicija, a grafički se mogu prikazati s pomoću hijerarhijskih struktura, sustava čvorova, poveznica itd. (Bessé <i>i sur.</i> 1997: 131)</i></p> <p>field: <i>filologija</i></p> <p>branch: <i>opće jezikoslovlje (lingvistika)</i></p> <p>project: <i>Hrvatsko jezikoslovno nazivlje</i></p> <p>subfield: <i>terminologija</i></p>	<p>domena <i>im. ž. r. (n. fem.)</i></p> <p>definition: <i>koherentno, stabilno i konvencionalizirano konceptualno područje na temelju kojega se interpretiraju određena značenja</i> (a coherent, stable and conventionalized conceptual area on the basis of which certain meanings are interpreted)</p> <p>source reference: Austin (2014)</p> <p>subordinate concept: <i>ciljna domena; domena varijantnosti; izvorna domena; primarna domena; sekundarna domena; temeljna domena</i></p> <p>English: domain</p> <p>German: <i>Domäne</i></p> <p>French: <i>domaine</i></p> <p>Russian: <i>домен</i></p> <p>note: <i>Primjerice značenja ljubavi, mržnje, straha itd. mogu se interpretirati samo u okviru domene emocija. Upravo stabilnost, koherentnost i konvencionaliziranost domene razlikuje od mentalnih prostora u teoriji konceptualne integracije, koji najčešće imaju lokalnu uporabu.</i></p> <p>field: <i>filologija</i></p> <p>branch: <i>opće jezikoslovlje (lingvistika)</i></p> <p>project: <i>Hrvatsko jezikoslovno nazivlje</i></p> <p>subfield: <i>kognitivna lingvistika</i></p>
--	---

Unnecessary entries within multiple domains/projects

As the term base *STRUNA* grew and the “big” domains like physics, mathematics, medicine, etc., were processed, many projects were involved covering more narrowly fo-

cused subdomains. As a result, experts are often not satisfied with the edits they find in *STRUNA*, which can be considered “theirs” and they do not want to share them with another project, that is, to put a like mark instead of

TABLE 9
AN EXAMPLE OF COINING A NEW TERM

etymology	dialectology
<i>konsonantizam riječi</i> <i>suglasnički sastav riječi</i> (the consonants of a word)	<i>konsonantizam</i> <i>inventar, realizacija, distribucija i podrijetlo suglasnika u kojemu jezičnom sustavu</i> (the consonant system of a language)
<i>vokalizam riječi</i> <i>samoglasnički sastav riječi</i> (the vowels of a word)	<i>vokalizam</i> <i>inventar, realizacija, distribucija i podrijetlo samoglasnika u kojemu jezičnom sustavu</i> (the vowel system of a language)

TABLE 10
TERMINOLOGICAL UNITS 'ETYMOLOGY' IN *STRUNA*

<i>etimologija</i> <i>im. ž. jd. (n. sg. fem.)</i> definition: <i>jezikoslovna disciplina koja istražuje podrijetlo, povijest i razvoj riječi</i> (a branch of linguistics concerned with the origin and historical development of a word) English: etymology German: <i>Etymologie</i> Russian: <i>этимология</i> field: <i>etimologija</i> project: <i>Hrvatsko jezikoslovno nazivlje</i>	<i>etimologija</i> <i>im. ž. jd. (n. sg. fem.)</i> definition: <i>podrijetlo, povijest i razvoj jezičnog oblika</i> (the origin and historical development of a linguistic form) English: etymology German: <i>Etymologie</i> Russian: <i>этимология</i> field: <i>etimologija</i> project: <i>Hrvatsko jezikoslovno nazivlje</i>
---	---

TABLE 11
TERMINOLOGICAL UNITS 'TAXONOMY' IN *STRUNA*

<i>taksonomija</i> <i>im. ž. jd. (n. sg. fem.)</i> definition: <i>sustav klasifikacije i imenovanja naziva</i> (classification and naming system) source reference: Bessé <i>i sur.</i> (1997) English: taxonomy njemački: <i>Taxonomie</i> French: <i>taxinomie</i> Russian: <i>таксономия</i> note: <i>U ovome je značenju naziv taksonomija blizak nazivu nomenklatura.</i> field: <i>filologija</i> branch: <i>opće jezikoslovlje (lingvistika)</i> subfield: <i>terminologija</i> project: <i>Hrvatsko jezikoslovno nazivlje</i>	<i>taksonomija</i> <i>im. ž. jd. (n. sg. fem.)</i> definition: <i>proučavanje i metodologija razredbe objekata i pojmova</i> (the scientific study and methodology of classifying objects and concepts) source reference: Bessé <i>i sur.</i> (1997.) English: taxonomy njemački: <i>Taxonomie</i> French: <i>taxinomie</i> Russian: <i>таксономия</i> field: <i>filologija</i> branch: <i>opće jezikoslovlje (lingvistika)</i> subfield: <i>terminologija</i> project: <i>Hrvatsko jezikoslovno nazivlje</i>
---	---

TABLE 12
TERMINOLOGICAL UNITS 'FUNGISTAT' IN *STRUNA*

<i>fungistatik</i> <i>im. m. jd. (n. sg. masc.)</i> definition: <i>tvar ili sredstvo koje zaustavlja rast gljivica</i> (substance or agent that stops the growth of fungi) allowed term: <i>mikostatik</i> English: fungistat German: <i>Fungistatikum</i> field: <i>temeljne medicinske znanosti</i> branch: <i>farmakologija</i> project: <i>Hrvatsko stomatološko nazivlje</i>	<i>fungistatik</i> <i>im. m. jd. (n. sg. masc.)</i> definition: <i>protugljivični lijek koji zaustavlja rast i razmnožavanje gljivica, ali ih ne ubija</i> (an antifungal drug that stops the growth and reproduction of fungi, but does not kill them) English: fungistat field: <i>temeljne medicinske znanosti</i> branch: <i>farmakologija</i> project: <i>Hrvatsko farmakološko nazivlje</i>
---	---

TABLE 13
TERMINOLOGICAL UNITS 'RETINA' IN *STRUNA*

<p>mrežnica <i>im. ž. jd. (n. sg. fem.)</i> definition: <i>dio unutarne očne ovojnice koji sadržava vidne stanice i koji može primiti svjetlosne podražaje</i> (the inner layer of the eye that contains visual cells and that can receive light stimuli) allowed term: <i>retina</i> English: <i>retina</i> Latin: <i>retina</i> note: <i>Mrežnica se dijeli na slijepi dio koji odgovara zrakastomu tijelu i šarenici te vidni dio koji odgovara žilnici.</i> field: <i>temeljne medicinske znanosti</i> branch: <i>anatomija</i> project: <i>Hrvatsko anatomsko i fiziološko nazivlje</i></p>	<p>mrežnica <i>im. ž. jd. (n. sg. fem.)</i> definition: <i>unutarnja očna ovojnica koja sadržava vidne stanice i pretvara svjetlosne podražaje u električne impulse koji se dalje živčanim vlaknima prenose u mozak</i> (the inner layer of the eye, which contains visual cells and converts light stimuli into electrical impulses that are further transmitted to the brain by nerve fibers) allowed term: <i>retina, unutarne očna ovojnica</i> English: <i>retina, inner layer of eyeball</i> Latin: <i>retina, tunica interna bulbi</i> note: <i>Mrežnica je građena od dva lista: neurosenzornoga sloja i mrežničnoga pigmentnog epitela, a odgovorna je za centralnu vidnu oštrinu, osjet za boje, periferni, fotopski, mezopski i skotopski vid. Ova dva lista prelaze jedan u drugi sprijeda na zjeničnome rubu, a straga u području glave vidnoga živca.</i> field: <i>kliničke medicinske znanosti</i> branch: <i>oftalmologija</i> project: <i>Hrvatsko oftalmološko nazivlje</i></p>
---	---

TABLE 14
TERMINOLOGICAL UNITS 'AUTOPSY' IN *STRUNA*

<p>obdukcija <i>im. ž. jd. (n. sg. fem.)</i> definition: <i>otvaranje i rastavljanje trupla radi ustanovljivanja uzroka smrti</i> (opening and dissecting the corpse to establish the cause of death) proposed term (IHJJ): <i>razudba</i> allowed term: <i>autopsija</i> deprecated term: <i>sekcija</i> English: <i>autopsy</i> German: <i>Obduktion, Sektion</i> Italian: <i>autopsia</i> field: <i>temeljne medicinske znanosti</i> project: <i>Hrvatsko stomatološko nazivlje</i></p>	<p>obdukcija <i>im. ž. jd. (n. sg. fem.)</i> definition: <i>pregled vanjskih i unutarnjih dijelova mrtvoga tijela radi utvrđivanja patologije organa, ozljeda te uzroka, načina, mehanizma i vremena smrti</i> (examination of the external and internal parts of the dead body in order to determine the pathology of the organs, injuries and the cause, method, mechanism and time of death) English: <i>autopsy</i> field: <i>kliničke medicinske znanosti</i> branch: <i>sudska medicina</i> project: <i>Forenzično-kriminalističko nazivlje</i></p>
--	---

redefining them. This is good for *STRUNA*'s end-user because a narrowly specialized expert will enter a new, more precise, or accurate definition. Still, on the other hand, it is the same concept that is defined differently, which is not in accordance with terminological norms, and the end-user does not know why the same concept appears twice.

The majority of such problems appear in medical terminology. In the Croatian Dental Terminology (2009 – 2010) project, concepts traditionally associated with basic medical sciences and clinical medical sciences were processed in addition to dental medicine concepts. The Croatian Anatomical and Physiological Terminology (2012 – 2013) covers the field of basic medical science; the Croatian Ophthalmological Terminology (2016 – 2018) the field of clinical medical science; the Croatian Pharmacological Terminology

(2016-2017) the field of basic medical science; and in the Forensic-Criminal Terminology (2018-2019) part of the concepts that belong to the fields of basic and clinical medical science are covered. Therefore, it is not surprising that there are many overlaps and disagreements.

For example, the Croatian Dental Terminology project has as many as 4,426 terminological units. Still, the editors covered too broad a spectrum and included concepts that the actual field experts would much better edit. For example, pharmacologists were bothered that a superordinate concept does not define the term *fungistat* (*fungistatik*). They also believe it is better and more logical that this terminological unit is included within pharmacology terminology instead of the dental one (Table 12).

Croatian Ophthalmological Terminology is the first clinical medical project in *STRUNA*, and the experts who worked as editors were not satisfied with some of the terminological units processed in the Croatian Dental Terminology and Croatian Anatomical and Physiological Terminology project. A joint decision was made that some terminological units would not be adopted into their project but defined anew¹⁰. Even though this decision resulted in a better and more precise definition for several ophthalmological concepts, the fact remains that the end-user is now faced with multiple results for the queries, for example, the term retina (Table 13).

Forensic sciences are transdisciplinary, including concepts from various, usually considered unrelated, domains. Consequently, within the project Forensic-Criminal Terminology, a certain amount of medical concepts was processed. The editors on the project were also dissatisfied with some definitions of concepts entered in the previous projects. An example of such a concept eventually entered anew is autopsy (Table 14).

Conclusion

Managing a big national term bank leads to a certain amount of turbidity in the front end of the search engine. *STRUNA* was conceived as a multidomain term base where field experts are crucial in planning and executing which terminological units will be processed. It was clear from the beginning that some end-users would find the results of their queries unsatisfactory. The terminological principles that have been the basis of terminology management in *STRUNA* are far from perfect, even when dealing with a single domain. Following them closely in a multidomain terminological database was often impossible. Even if we put aside the needs of the end-users, cases of multiple entries, such as the ones we presented in the previous chapters, represent a problem from a terminological and data management point of view. Ambiguity needs to be resolved if one is planning semi-automatic data management, which, considering current trends in IT and modern linguistics, is a mandatory characteristic of any language tool.

Unfortunately, as we have shown, a certain amount of sidestepping the one-term-one-concept principle is unavoidable, even when working with terminological units within a relatively small domain. When processing termi-

nological units from multiple domains that will be simultaneously searched based on the end-user query, additional steps to optimize the search engine usability will have to be taken. Even though the overarching purpose of *STRUNA* is the standardization of Croatian specialized languages, we must keep in mind that the end-user's satisfaction must be, if not our primary goal, then at least one of the fundamental ones. Studies have shown that end-users who have repeated anecdotal bad experiences with search engines tend to characterize the engine in question as unreliable, despite the general quality of the results¹⁴.

A solution to this problem will not be easily reached since multiple aspects of terminology management must be reevaluated and tested before a final solution is implemented. One of the things that we would like to avoid is implementing additional steps to query submission, for example, forcing the end-user to narrow the search span by choosing which domains and subdomains to include in the query. The main reason is that we would like to save time and energy spent waiting for the result to be presented on the end-user's screen.

With the rapid development of AI-based software solutions, one possible solution is implementing an analogical reasoning module based on human-in-the-loop AI, which could filter the search results based on the real needs of the end-user. Such solutions have shown to be useful in bridging the gap between surface matches and simple keyword searches and the actual abstraction of relevant data¹⁵. Furthermore, we are currently testing possibilities of implementing a "terminological unit generator" driven by OpenAI's GPT-4¹⁶, a language model trained on a large corpus of text. GPT-4 can generate high-quality definitions for various terms and concepts. It can "understand" natural language and provide accurate and informative definitions contextualized within their respective fields or domains. Additionally, it can be customized to incorporate specific terminology and jargon relevant to a particular industry or domain, making it a valuable tool for expanding the services a terminological database like *STRUNA* is offering to its end-users.

We aim to bring *STRUNA* closer to the current data retrieval standards while increasing our end-user's satisfaction to the point where they feel confident that we are providing a valuable service and retrieving accurate and relevant information.

REFERENCES

1. HUDEČEK L, MIHALJEVIĆ M, Terminološki priručnik (Institut za hrvatski jezik i jezikoslovlje, 2009). — 2. BRATANIĆ M, OSTROŠKI ANIĆ A, Konceptija i ustrojstvo terminološke baze *STRUNA*. In BRATANIĆ M, BRAČI I, PRITCHARD B (Eds) *Od Šuleka do Schengena: Terminološki, terminografski i prijevodni aspekti jezika struke* (Institut za hrvatski jezik i jezikoslovlje, 2015). — 3. BERGOVEC M, RUNJAIĆ S, Harmonization of Multiple Entries in the Terminology Database *STRUNA* (Croatian Special Field Terminology) In: Proceedings (10th Terminology and Knowledge Engineering Conference, Madrid, 2012). — 4. CHIN JP,

- DIEHL VA, NORMAN KL, Development of an instrument measuring user satisfaction of the human-computer interface. In: Proceedings (Conference on Human Factors in Computing Systems, Washington, 1988). — 5. MAHMOOD, MA, BECKER JD, Information Systems, 2 (1985) 37. — 6. WÜSTER E. Einführung in die allgemeine Terminologielehre und terminologische Lexikographie. (Springer, Wien/New York, 1979). — 7. NAHOD B. Od polisemna leksema do homonimna naziva. In: Proceedings (Zbornik Hrvatskog društva za primjenjenu lingvistiku Osijek, 2011). — 8. MIHALJEVIĆ M, HUDEČEK L, JOZIĆ Ž (Eds) *Hrvatsko jezikoslovno*

nazivlje. (Institut za hrvatski jezik i jezikoslovlje, Zagreb 2022). — 9. BAREŠIĆ A et al. Hrvatsko antropološko nazivlje (Institut za hrvatski jezik i jezikoslovlje, Zagreb, 2015). — 10. NAHOD B, VUKŠA NAHOD P, BJELOŠ M, Jazykovedny Casopis 2 (2017) 86. — 11. BARTOLEC BLA-GUS G. Pragmalingvističko nazivlje. In: MIHALJEVIĆ M, HUDEČEK L, JOZIĆ Ž (Eds) Hrvatsko jezikoslovno nazivlje. (Institut za hrvatski jezik i jezikoslovlje, Zagreb 2022). — 12. BIRTIĆ M, BRAČ I. Dopuna i dodatak u različitim jezikoslovnim teorijama. In: MIHALJEVIĆ M,

HUDEČEK L, JOZIĆ Ž (Eds) Hrvatsko jezikoslovno nazivlje. (Institut za hrvatski jezik i jezikoslovlje, Zagreb 2022). — 13. IVŠIĆ MAJIĆ D, Etimološko nazivlje. In: MIHALJEVIĆ M, HUDEČEK L, JOZIĆ Ž (Eds) Hrvatsko jezikoslovno nazivlje. (Institut za hrvatski jezik i jezikoslovlje, Zagreb 2022). — 14. CHEN L, IEEE Access, 8 (2020) 79024. — 15. KANG H, Qian X, Hope T, Shahaf D, Chan J, Kittur A, ACM Transactions on Computer-Human Interaction 1 (2022). doi:10.1145/3530013. — 16. OpenAI GPT-4. <https://openai.com/product/gpt-4>.

B. Nahod

Institute of Croatian Language and Linguistics, Ulica Republike Austrije 16, 10000 Zagreb, Croatia

e-mail: bnahod@ihj.hr

NAZIVI KOJE DIJELIMO ILI NE ŽELIMO DIJELITI S DRUGIMA – O VIŠESTRUKIM UNOSIMA U STRUNI

SAŽETAK

Tijekom dugogodišnjega rada na obradi hrvatskoga strukovnog nazivlja u STRUNI susreli smo se s dvama problemima koji za posljedicu imaju zbunjujući učinak na krajnjega korisnika. S jedne strane, riječ je o homonimiji. Taj je problem zapravo očekivan jer jedan naziv može imati više različitih značenja unutar više struka ili unutar jedne struke. To je u STRUNI riješeno tako da je omogućen dvostruki ili višestruki unos. S druge strane, katkad se događa da je isti naziv definiran dva puta. Takvi su se slučajevi od početka rada na STRUNI pokušali izbjeći uvođenjem opcije pridruživanja postojećega naziva, no obrađivači rijetko to čine. Navedeni se problemi mogu i moraju riješiti u budućnosti. U radu ćemo donijeti kako korisnici pretraživanjem STRUNE vide takve primjere te dosadašnje primjere iz STRUNE koji pokazuju rješavanje takvih problema unutar pojedine struke, ali i nakon obrade naziva više struka.

