



A national term bank as an instrument of terminology planning

Prof. Maja Bratanić, coordinator  
Institute of Croatian Language and Linguistics  
IULA, Barcelona, February 4, 2014

A national term bank as an instrument of terminology planning (the case of Croatia(n))

---

☞ Radoslav Katičić: “in southeastern Europe, typically yet paradoxically, an unusual phenomenon occurs: *new languages* have a rich history”

lack of language and terminology policy

---

- ☞ Croatian was not used as official language till mid 19th; terminology didn't develop
  - ☞ Latin: until 1848, German: 1850 – 1868, Hungarian as obligatory subject in 1827
- ☞ Bogoslav Šulek – “father of Croatian scientific terminology” (purism, neologisms)
- ☞ guiding principle: “avoid extremes, especially excessive purism and unnecessary classicism”
- ☞ 20th century – centralization and forceful unification
- ☞ since 1991 - language planning still rather inarticulate
- ☞ terminology planning left very much aside

initiation of Struna

---

- ☞ development of Croatian Special Field Terminology - STRUNA
- ☞ initiated in 2007 by Council for Standard Croatian Language (dismissed in 2012)
- ☞ financed by Croatian Science Foundation
- ☞ Institute of Croatian Language and Linguistics (IHJJ) serves as national coordinator

Coordination project – 1st phase

---

- ☞ in 2008 - a Coordination project launched
- ☞ Department for the Croatian Standard language of the Institute
- ☞ first version of the **database** was lexicographically oriented
- ☞ since then altered profoundly (currently version 5 in preparation)
- ☞ *Croatian Manual of Terminology*
- ☞ initial workshops for field experts

education & termbank development

---

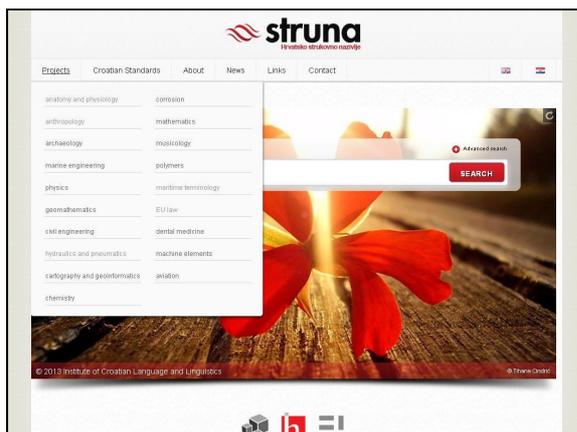
- ☞ 2009 first partner projects (20 so far)
- ☞ Struna transferred to the Dept. of General Linguistics (activity of IHJJ, not separate body)
- ☞ education of terminologists started
- ☞ Infoterm TM courses; IULA courses
- ☞ ISO standards; Unesco Guidelines
- ☞ fundamental literature
- ☞ existing termbanks and termbases

## aims of Struna

- necessary prerequisites for the creation and standardization of the Croatian terminology across various domains
- via coordination of domain experts and
- terminologists/terminographers and language experts
- establish framework for a national terminology policy
- foundations for more structured education in terminology
- development of terminology infrastructure

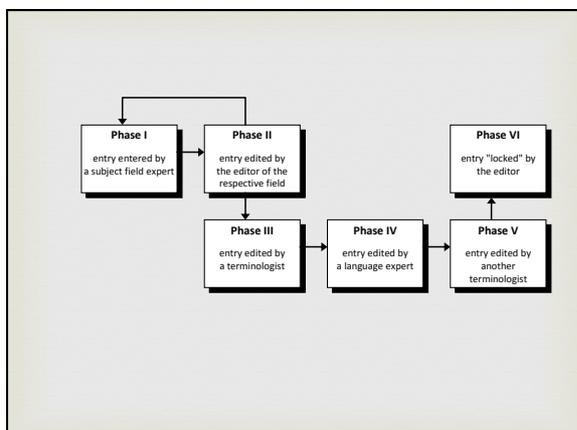
## selection of the projects

- ☞ public tender for LSP projects (2009-2011)
- ☞ heads of projects from university and research
- ☞ budget – € 263,300 so far (allotted to partner projects)
- ☞ ISO 15188 – 1400 terms
- ☞ 44,000 terms
- ☞ 30,000 entries (including synonyms)
- ☞ no new projects since the end of 2011
- ☞ Struna was opened for public on the web in February 2012



## model of terminology work

- ☞ collaborative terminology work between subject field experts and language experts
- ☞ shared responsibility
- ☞ IHJJ provides terminologists and Croatian language editors
- ☞ terminographers (doctoral and postdoctoral students, none work on Struna full time)
- ☞ workflow divided into six phases



## challenges and misconceptions

- ☞ expectation that all compilers actually accept the idea of *establishing* Croatian LSP terminology which implies:
- ☞ providing Croatian replacements for internationalisms instead of merely authorizing the prevailing usage
- ☞ distinguishing encyclopedic vs. lexicographic vs. terminological definition
- ☞ tendency of subject field experts to include encyclopedic information (→ *context* and *note*)

## advantages...

- doing terminology in a national term base (as opposed to professional or domain TB):
  - uniform terminographic description
  - cooperation between domain experts and terminologists/language experts
  - faster term validation

## ...and drawbacks

- concept overlapping is unavoidable
- domain discourse variation (terms of Latin and Greek origin in medical terminology; adjectives in biosciences ...)
- large number of people working in the term base → coordination among projects increasingly difficult
- compromises necessary

## users

- another disadvantage of a relative independence of the domain experts: different projects may address different categories of users
- translator's needs vs. general public vs. field experts
- for the model to achieve its full potential: higher degree of coordination among experts and terminographers and a more rigid adherence to the workflow required

## underlying methodology

- normative termbase
- largely organized according to the principles of the GTT and relevant ISO/TC 37 terminological standards
- TBX compatible

## working interface

The screenshot shows the e-Struna web application interface. On the left is a vertical navigation menu with options like 'Početna', 'Nazivi', 'Upute', 'Pregled', 'Popisnica', 'Prijavi', 'Statistika', 'Aktivnost na projektu', 'Upravljaj nazivima', 'Upravljaj terminima', and 'Komentari'. The main content area includes a header with the e-Struna logo and a navigation bar. Below the header, there are several sections: 'Informacije' with project details, 'Statistika' with numerical data, and 'Aktivnost na projektu' with a list of recent activities.

## working interface

This screenshot shows a detailed view of the 'Popis naziva' (List of terms) in the e-Struna interface. It features a table with columns for '%', 'naziv', 'graj-ekst', 'porijeklo', 'grana', 'postojecnost', and 'redoslijed'. The table lists various terms such as 'propiliona-epigalaktol', 'alfa-epigalaktol', 'peptidna integracija', 'hidrolitička hidroliza', 'jelena', 'epigalen', 'elektromotorna funkcija', 'glavna molekularna funkcija', 'glavni molekularni funkcije', 'glavni strukturni funkcije', 'kratkostrešni peptid', 'kardiolipin', 'dihaloalkilna polihidroks', and 'koordinativna'. Each row includes a percentage value (15%), a term name, and several status indicators (yes/no) for different criteria.

## data categories

- 46 categories altogether
- administrative data categories recorded automatically
- **concept level:** subject field and subfield, definition, subordinate concepts, sources, interdisciplinary term label, note, formula, equation, symbol, picture, shared resources
- **term level:** preferred term, synonyms, context, neologism label, equivalents, abbreviations

## data categories ctd.

- **language level:** grammatical information, foreign term label
- **mandatory categories:**
  - preferred term
  - grammatical information
  - definition
  - equivalent in English
- **synonyms:** admitted, deprecated, obsolete, colloquial and proposed

### individual entry

### flexible presentation: one or two tabs

## mandatory categories

## definition

- ☞ must reflect the conceptual structure of a certain domain
- ☞ degree of variation in definition structure
- ☞ different types of definitions used in e.g. biomedical domain, engineering, anthropology or EU law

## synonyms

- ☞ labeled according to their normative status
- ☞ recommended term
- ☞ admitted term
- ☞ deprecated term
- ☞ obsolete term
- ☞ colloquial term
- ☞ *proposed term* – if subject field experts do not accept the language experts' solution

vozna staza	
<b>definicija</b>	staza na aerodromu na zemlji za vožnju zrakoplova koja osigurava vezu između jednoga i drugoga dijela aerodroma
<b>istoznačnice</b>	<b>deputetni naziv:</b> staza za vožnju, voznica <b>zastarjeli naziv:</b> ruina staza <b>žargonizam:</b> rulnica
<b>istovjednice</b>	<b>engleski:</b> taxiway <b>njemački:</b> Rollweg, Rollbahn <b>francuski:</b> voie de circulation
<b>kratice</b>	<b>engleska:</b> TWY
<b>podređeni nazivi</b>	brza izlazna vozna staza, vozna staza do pozicije, vozna staza na stajanci
<b>razredba</b>	<b>polje:</b> tehnologija prometa i transport <b>grana:</b> zračni promet <b>projekat:</b> Usustavljanje i normiranje temeljnoga hrvatskoga zrakoplovnog nazivlja

treći kutnjak	
<b>definicija</b>	kutnjak smješten u bočnome dijelu gornjega ili donjega zubnog luka iza drugoga kutnjaka, obilježen je širokom griznom plohom, služi za mljevenje hrane
<b>radna definicija</b>	kutnjak smješten u bočnome dijelu gornjega ili donjega zubnog luka iza drugoga kutnjaka, obilježen je širokom griznom plohom, služi za mljevenje hrane
<b>istoznačnice</b>	<b>deputetni naziv:</b> treći molar, umnjak <b>žargonizam:</b> osmica, zub mudrosti
<b>istovjednice</b>	<b>engleski:</b> cranter, wisdom tooth, third molar <b>njemački:</b> Achter, Weisheitszahn, dritter Malzahn <b>italijanski:</b> dente molare tertius <b>talijanski:</b> terzo molare
<b>podređeni nazivi</b>	gornji treći kutnjak, donji treći kutnjak
<b>razredba</b>	<b>polje:</b> temeljne medicinske znanosti <b>grana:</b> anatomija <b>projekat:</b> Hrvatsko stomatološko nazivlje

karta <small>im. 2. gl.</small>	
<b>definicija</b>	prikaz prostora s pomoću znakova koji omogućuju utvrđivanje odnosa među elementima toga prostora
<b>predloženi naziv (HLLJ)</b>	zemljovid
<b>istoznačnice</b>	<b>zastarjeli naziv:</b> mapa, tabula
<b>istovjednice</b>	<b>engleski:</b> map <b>njemački:</b> Karte <b>francuski:</b> carte
<b>podređeni nazivi</b>	relna karta, tematska karta, topografska karta
<b>napomena</b>	Elementi prostora jesu objekti, područja ili teme. Karta je najbolje prikaz površine nebeskoga tijela ili svoda. Razlikuju se topografske i tematske karte te realne i virtualne karte; ne postoje druge podjele karata. Prema mnogo obitno prikazuju geoprostor, karte mogu prikazivati bilo koji prostor, svami ili zamisljeni, bez obzira na sadržaj ili mjerilo, primjenka karte mozga, karte DNA-a ili karte svemira.
<b>razredba</b>	<b>polje:</b> geodezija <b>grana:</b> kartografija <b>projekat:</b> Kartografsko-nazivljeno nazivlje

## contextual information

- context traditionally used as a source of linguistic information
- defining context (ISO 704)
  - ☐ context as a source of domain knowledge information on the concept
- knowledge-rich contexts
- context and note used as a source of relevant linguistic and extralinguistic information

## contextual information ctd.

- context always taken from a source
- note usually written by field specialists or, less commonly, terminologist
- contextual information (context and note)
  - ☐ concept level (conceptual relations among concepts in the framework or entire domain; multidimensionality of a concept)
  - ☐ term level (usage information)
  - ☐ encyclopedic information

## sources of conceptual relations

- ☞ **Broca's area** – (*definition*) the part of left frontal lobe concerned with the production of speech
- ☞ *context* – Pierre Paul Broca (1824 – 1880), neurologist, claimed in 1861 that an aphasia in a certain patient (whose brain he examined post mortem) had been the result of an accurately circumscribed lesion of the posterior third of the second and third frontal convolutions. This specific area was later named Broca's area.

## ☞ Broca's area

- ☞ *note* – Recent research deals with the broader role of this region of the brain as an interface between action and perception, the central role of which is the coordination of perception and motor functions that serve as the basis of verbal and non-verbal communication.

## concept's multidimensionality

### ☞ index

- ☞ (*definition*) a sign that refers to the signified based on their spatial, temporal or causal relations
- ☞ (*context*) (...) index can be understood as a consequence of a certain cause (e.g., in the case of a symptom of an illness), as a sign of some possibility (e.g., clouds are a sign of the possibility of rain), as a trace left by a perpetrator, as a public response to an event, etc.

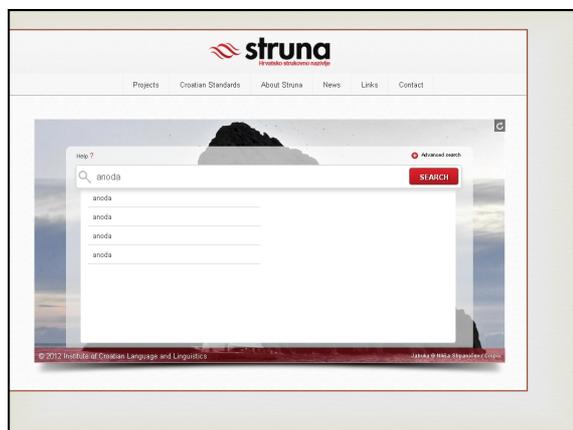
## term level information

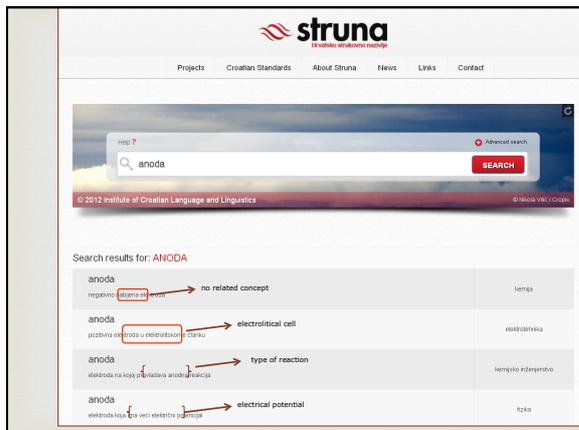
### ☞ ethnosemantics

- ☞ (*note*) The terms *ethnosemantics*, *ethnoscience*, or *new ethnography* are synonyms for an approach to the study of culture as a body of knowledge that developed in the 1960s (...).

## harmonization

- multiple records of equivalent terms with identical (or roughly identical) meanings but different definitions
  - general language concepts in LSP
  - *interdisciplinary term* label introduced to avoid multiple records of the same concept
  - analysis of multiple terminological entries carried out
- ☞ (Bergovec, M.; Runjaić, S. Harmonization of Multiple Entries in the Terminology Database Struna (Croatian Special Field Terminology). // *Proceedings of the 10<sup>th</sup> Terminology and Knowledge Engineering Conference (TKE 2012)*, 19–22 June 2012, Madrid, Spain / ur. Guadalupe Aguado de Cea et al. str. 231–241.)





## harmonization ...

- 80 percent of multiple entries turned out to be conceptual synonyms
- related domains (e. g. polymers, corrosion, and chemistry) defined the same concepts from a different perspective
- special field autonomy preserved to a reasonable level (*heart*)
- mechanisms introduced in the term base to prevent uncontrolled multiplication



## future of Struna

- including reliable existing terminological resources
- self-funded projects
- individual initiatives (authors)
- projects carried out internally at IHJJ

