

# Monolingual Lexicography

## Communication in Slovene Project

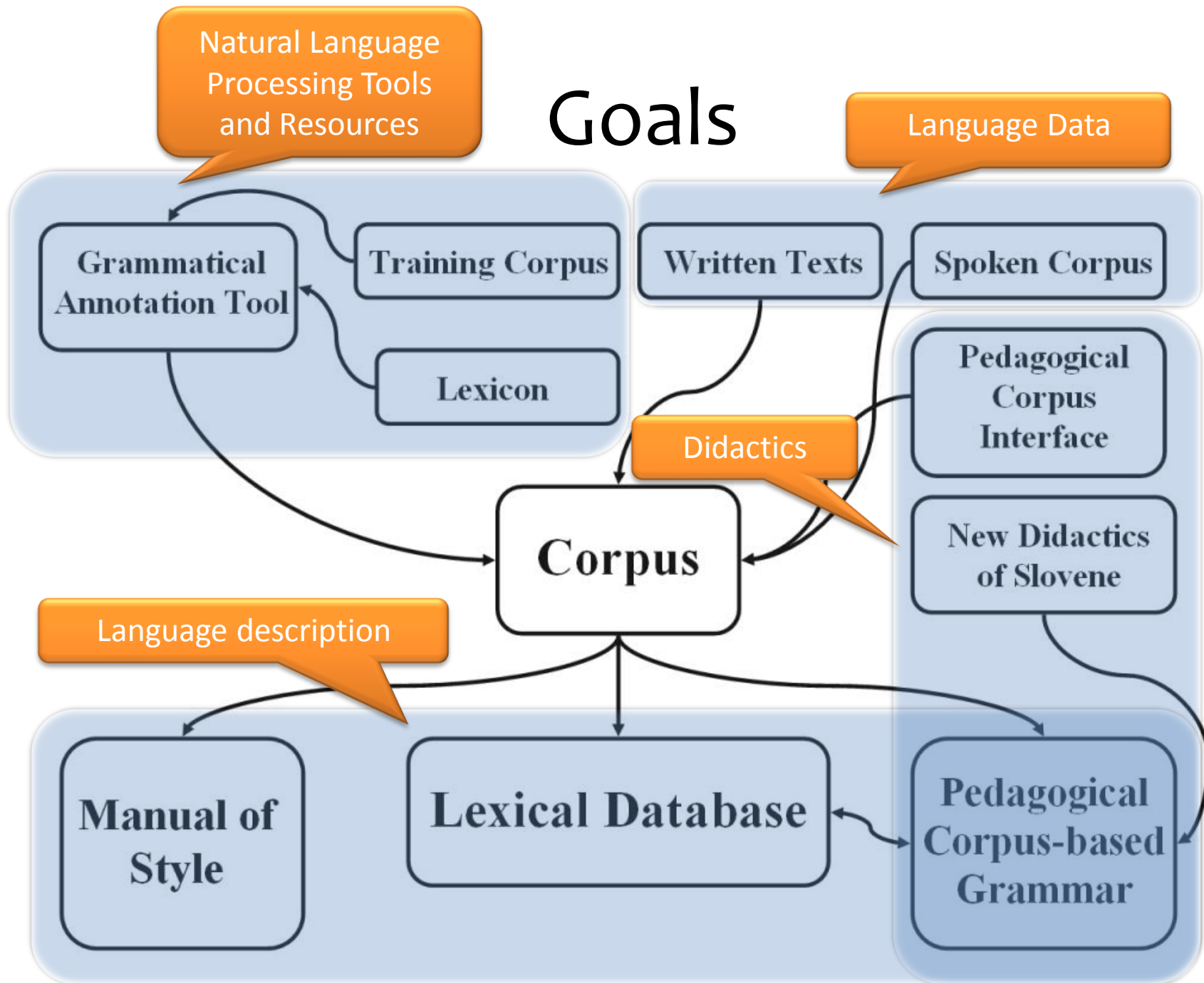
### Slovene Lexical Database

Session 2

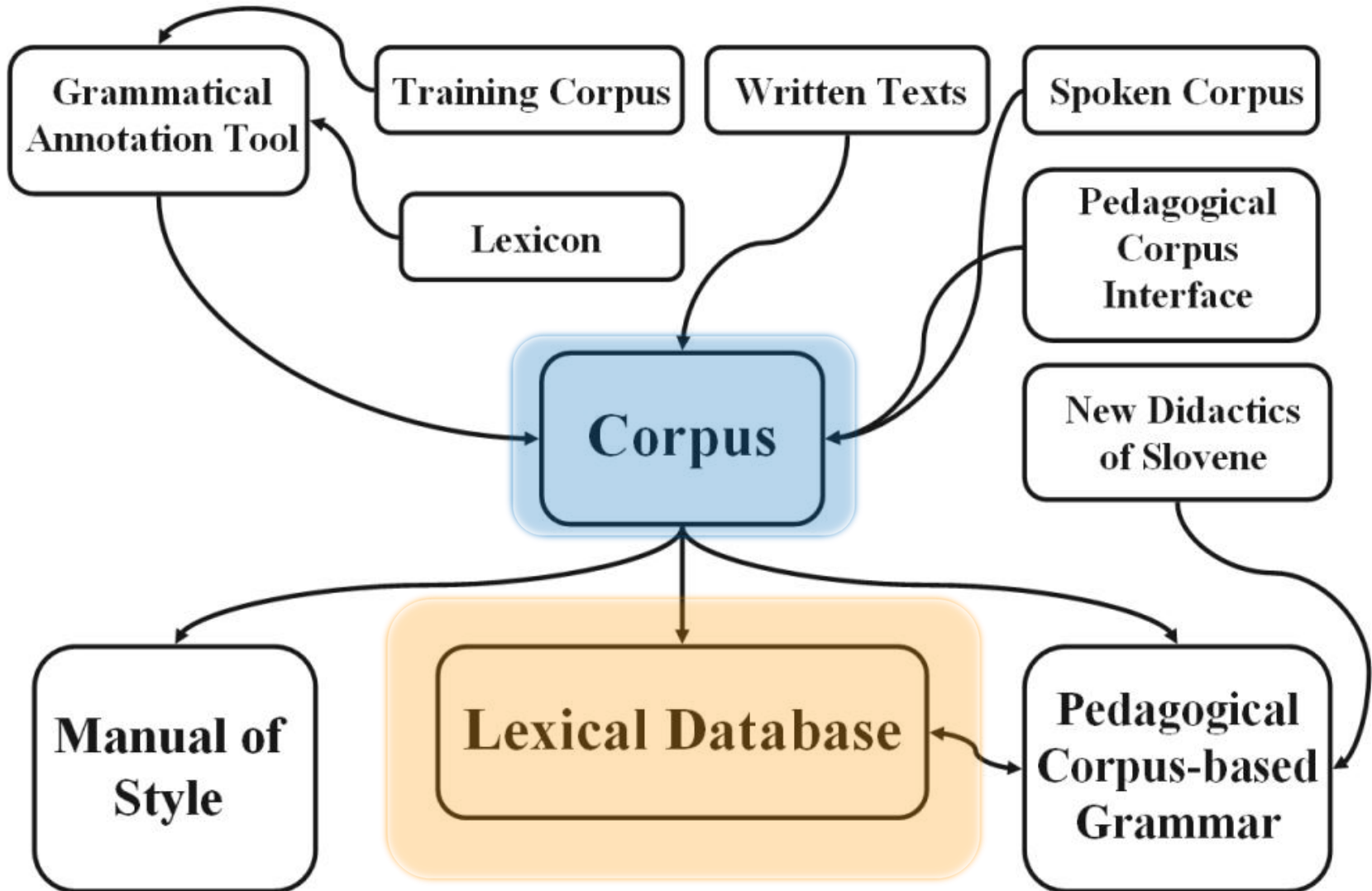
# “Communication in Slovene”

- <http://www.slovenscina.eu>
- European Social Fund + Ministry of Education
- Leading partner: Amebis, d. o. o., Kamnik
- Duration: June 2008 - December 2013
- Total value: 3.2 million Euro
- Project consortium:
  - [Amebis, d. o. o., Kamnik](#)
  - [Jozef Stefan Institute](#)
  - [University of Ljubljana](#)
  - [Scientific Research Centre of the Slovenian Academy of Sciences and Arts](#)
  - [Trojina, Institute for Applied Slovene Studies](#)

# Goals



# Goals



# Slovene Lexical Database - Timeline

- Number of lexical units: minimum 2,500
- June-October 2008: preparation
- November 2008-June 2009: specifications



# Legal aspects

- Creative Commons

- Attribution
- Share Alike
- Noncommercial



- Availability

- On-line (<http://www.termania.net/>)
- Dataset (<http://www.slovenscina.eu/>)

- Owner: Ministry of Education and Sports

# Past experience

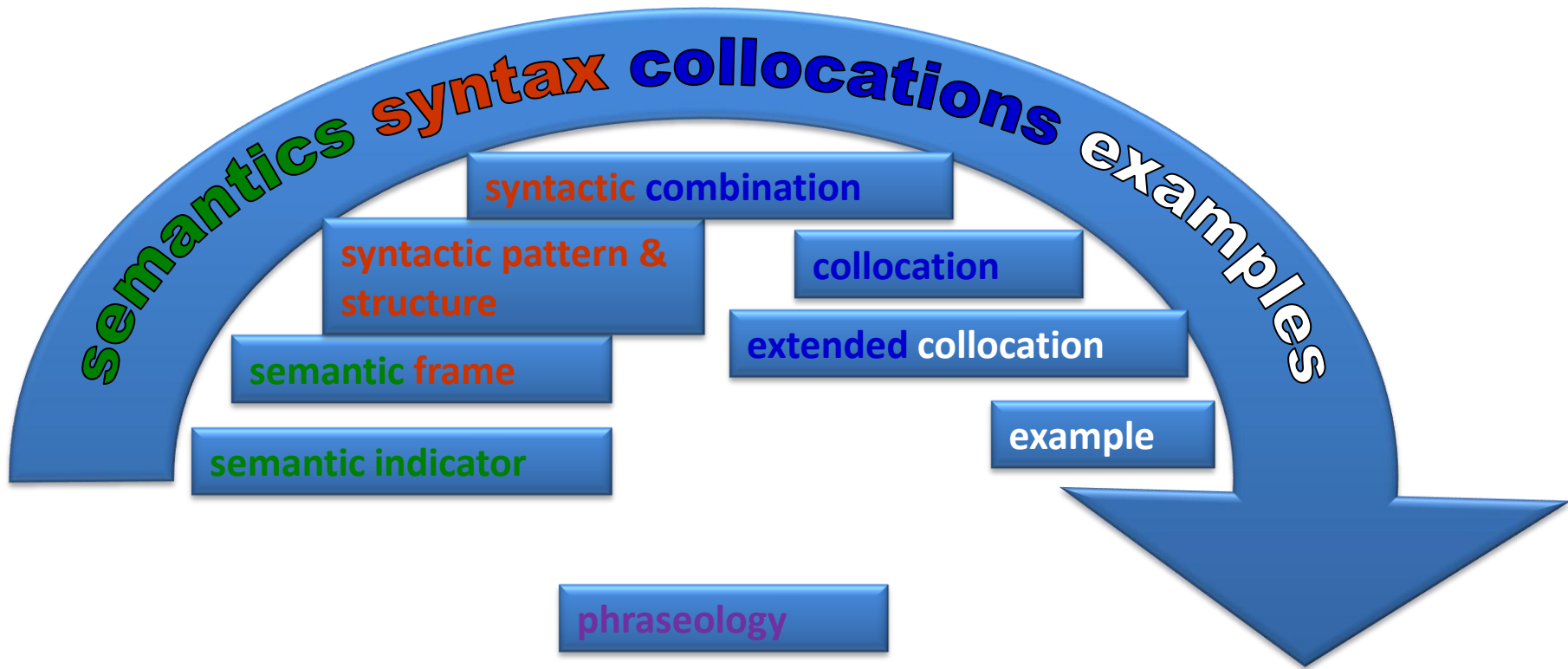
- **International (early):**
  - GENELEX (1990-94)
    - LE PAROLE (1993-98)
    - SIMPLE (1998-2002)

---
  - ACQUILEX I, II (- 1995)
  - ILC- DELIS ...
- **Individual languages:** [elexico](#) (DE), [CLIPS](#) (IT), [CORNETTO](#) (NL), [ALFALEX](#) (FR), [STO](#) (DK), [ADESSE](#) (SP), [GRIAL](#) (SP), [CEGLEX](#) (PL), [SALDO](#) (S), [BLF](#) (FR), PRALED (CZ), ...
- **Important for us:** FrameNet, Corpus Pattern Analysis, DANTE, COBUILD

# Basics

- corpus data analysis
- lexicogrammatical approach
  - semantics and syntax are not separated
  - valency – colligation – collocation
- meaning = meaning potential
  - is not stable (norms & exploitations)
- lumpers vs. splitters = splitters
- lexicography first, NLP second





## I. LEXICAL UNIT

- **headword** *to squeeze*
- **part-of-speech** verb

## II. SENSE

- **indicator** **1. grip firmly**  
If a PERSON squeezes an OBJECT, s|he presses it firmly, usually with his|her hands.
- **frame** **2. press out liquid**  
If a PERSON squeezes a LIQUID or a SOFT SUBSTANCE out of an OBJECT, s|he gets the liquid or substance out by pressing the object.

## III. SYNTAX

- **structure** **vb-obj**
- **pattern** *sb squeezes sth*

- **combination** (to squeeze your eyes shut)

## IV. COLLOC'S

- **collocation** to squeeze (sb's) [hand, arm]

## V. EXAMPLES

- **example** *I squeezed her hand gratefully.*

*She immediately squeezed the poison out and that probably saved her life.*

- **multi-word unit** (only nouns and adjectives)

## VI. PHRASEOLOGY

- **phraseological unit** to squeeze a quart into a pint pot

# I. Lexical Unit

- link to the lexicon
  - morphosyntactic information
  - corpus frequency
  - pronunciation etc.
- additional grammatical information
  - can be inferred (un/countability etc.)
  - manual (part-of-speech subtypes etc.)

## II. Semantic Level

- Semantic Indicators
  - simple EFL-like explanations or synonyms forming a sense menu
  - self-explanatory in relation to each other
- Semantic Frames
  - COBUILD / FrameNet / Corpus Pattern Analysis
  - combination of the systems

# Semantic Indicators - koža

**koža** *samostalnik*

## **1. vrhnji del telesa**

**1.1 pri človeku**

**1.2 pri živali**

## **2. odstranjen vrhnji del živalskega telesa**

## **3. ovoj ali lupina**

# Semantic Indicators – svitati se

**svitati se** *verb*

**1 delati se dan; daniti se**

**2 dojemati**

**2.1 nekaj vedeti**

**3 napovedovati se**

# Semantic Frames

- identification of verb/semantic arguments
  - prototypical pattern – “the norm” (Hanks)
  - the headword in its syntactic environment
- identification of semantic types in particular syntactic positions
- the semantic scenario
  - a full-sentence definition making a link between the arguments and the situation (FN) typical for a particular sense

# Semantic Frame

## 2. dojemati

### 2.1 nekaj vedeti

če se **ČLOVEKU** svita o nekem **DEJSTVU**, potem o tem nekaj ve ali sluti



# III. Syntactic Level

- **semantic frame** (between semantics and syntax)
  - semantic arguments in capital letters (ID-ed)
  - linked with collocates via syntax
- **syntactic structures** (formal)
  - clause and phrase level (all POS; only for NLP)
  - the number of syntactic structures is finite (SLB ~290)
  - source: word sketches (Sketch Engine)
- **syntactic patterns** (verbalized)
  - valency (only verbs; for lexicography and NLP)
- **syntactic combinations**
  - more than basic patterns: "to squeeze your eyes shut"

# Syntactic Structures – koža

## 4 vrhnji del telesa

### 1.1 pri človeku

- pbzo SBZ0 [občutljiva, suha, mastna] koža
- SBZ0 sbz2 koža [obraza, telesa, rok, lasišča]
- SBZ0 pod sbz6 koža pod [pazduho, očmi]
- gbz SBZ4 [dražiti, pomirjati, hladiti] kožo

# Syntactic Patterns – svitati se

## 2. dojemati

### 2.1 nekaj vedeti

če se **ČLOVEKU** svita o nekem **DEJSTVU**, potem o tem nekaj ve ali sluti

- komu se svita se o čem
- komu se svita kaj

# IV. Collocation Level

- **SEMANTIC FRAME:**

če se ČLOVEKU svita o nekem DEJSTVU, potem o tem nekaj ve ali sluti

- **SYNTACTIC STRUCTURES AND PATTERNS:**

NOUN – koža

pbzo SBZo

SBZo sbz2

SBZo pod sbz6

gbz SBZ4

VERB – svitati se

komu se svita se o čem

komu se svita kaj

If a part of syntactic patterns are collocational, they are shown on the collocation level.

- **COLLOCATIONS**

- [občutljiva, suha, mastna] koža
- koža [obraza, telesa, rok, lasišča]
- koža pod [pazduho, očmi]
- [dražiti, pomirjati, hladiti] kožo

# I. Examples

- **COLLOCATIONS**

- [občutljiva, suha, mastna] **koža**
- **koža** [obraza, telesa, rok, lasišča]
- **koža** pod [pazduho, očmi]
- [dražiti, pomirjati, hladiti] **kožo**

- **EXAMPLES (TBL + GDEX)**

- *Tonik je namenjen je **občutljivi koži** in ne vsebuje alkohola.*
- ***Koža rok** postane pozimi občutljivejša.*
- *Opažate na **koži pod očmi** prezgodnja znamenja staranja?*
  
- *Se vam že kaj **svita**, o čem govorim?*
- *Petru pa se pričinja **svitati** o nekdanji zvezi med Chandlerjem in Heather.*
- *Holly je na svojem stolu v klubu Diva zastokala in se prijela za glavo, ko se ji je začelo **svitati**, kaj se bo zgodilo.*

FOR  
WHAT

reference

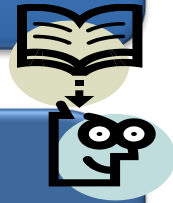


FOR  
WHOM

general user  
school population  
Slovene as foreign language

WHAT

semantic info  
menus + frames  
collocations  
corpus examples



natural language  
processing

computer  
linguist



semantic frames  
syntactic structures  
syntactic patterns  
other grammatical info



# Corpus Data & Authoring Tools

- FidaPLUS – Gigafida
- Sketch Engine: [www.sketchengine.co.uk](http://www.sketchengine.co.uk)
  - Slovene Sketch Grammar
  - Tickbox Lexicography
  - GDEX
- DWS
  - IDM Dictionary Production System
  - iLex

# TBL – Word Sketch (to love)

user: Simon Krek corpus: Fida PLUS 620m (SLD sketch grammar)

[Concordance](#)  
[Word List](#)  
[Word Sketch](#)  
[Thesaurus](#)  
[Sketch-Diff](#)  
[? Help on main menu](#)

[Save](#)  
[Change options](#)  
[Turn on clustering](#)  
[More data](#)  
[Less data](#)  
[Switch menu position](#)

**ljubiti** (*glagol*) Fida PLUS 620m (SLD sketch grammar) freq = 39356

kako-kdaj?	7198	5.3
<input checked="" type="checkbox"/> strastno	134	64.24
<input checked="" type="checkbox"/> neizmerno	150	63.12
<input checked="" type="checkbox"/> neskončno	93	51.73
<input type="checkbox"/> nesmrtno	36	51.49
<input type="checkbox"/> znova	290	49.92
<input type="checkbox"/> vedno	562	47.95
<input type="checkbox"/> nad vse	145	47.83
<input type="checkbox"/> brezpogojno	52	47.35
<input type="checkbox"/> iskreno	83	45.13
<input type="checkbox"/> noro	40	42.6
<input type="checkbox"/> ponovno	157	41.24
<input type="checkbox"/> srčno	41	40.47
<input type="checkbox"/> zelo	431	40.35
<input type="checkbox"/> tako	566	38.53
<input type="checkbox"/> resnično	58	37.63
<input type="checkbox"/> preprosto	95	36.85
<input type="checkbox"/> brezmejno	18	36.81
<input type="checkbox"/> bolj	309	36.52

veznik	1903	1.6
<input type="checkbox"/> in	221	43.18
<input type="checkbox"/> a	131	38.35
<input type="checkbox"/> vendar	130	37.95
<input type="checkbox"/> čeprav	70	34.25
<input type="checkbox"/> ker	107	33.38
<input type="checkbox"/> toda	56	32.63
<input type="checkbox"/> če	101	31.75
<input type="checkbox"/> zato	73	30.73
<input type="checkbox"/> kakor	34	30.44
<input type="checkbox"/> da	299	28.95
<input type="checkbox"/> ali	54	28.58
<input type="checkbox"/> ampak	45	27.18
<input type="checkbox"/> ko	81	26.6
<input type="checkbox"/> kar	50	23.55
<input type="checkbox"/> kadar	16	22.89
<input type="checkbox"/> dokler	17	22.12
<input type="checkbox"/> kajti	20	21.81
<input type="checkbox"/> saj	50	20.98

predlog	134	0.0
<input type="checkbox"/> kot	46	26.26
<input type="checkbox"/> od	23	18.13

pred-za	1874	0.5
<input type="checkbox"/> z	756	32.52
<input type="checkbox"/> kot	122	21.28
<input type="checkbox"/> zaradi	71	20.56
<input type="checkbox"/> brez	39	17.48
<input type="checkbox"/> ob	55	15.15
<input type="checkbox"/> do	70	15.08
<input type="checkbox"/> iz	74	14.66
<input type="checkbox"/> med	50	13.82
<input type="checkbox"/> v	268	13.65
<input type="checkbox"/> na	170	13.0
<input type="checkbox"/> od	39	9.82
<input type="checkbox"/> po	35	6.78
<input type="checkbox"/> pri	23	5.95

z-d	640	1.7
<input type="checkbox"/> srce	87	45.79
<input type="checkbox"/> ljubezen	42	34.13
<input type="checkbox"/> ženska	49	31.67
<input type="checkbox"/> maček	16	31.55
<input type="checkbox"/> žena	19	24.53
<input type="checkbox"/> moški	21	23.41
<input type="checkbox"/> moč	15	20.0

po-d	75	0.6
<input type="checkbox"/> smrt	18	32.34

kot-d	135	3.1
<input type="checkbox"/> Slovenec	17	29.5

v rodil	1624	10.4
<input type="checkbox"/> nada	238	73.21
<input type="checkbox"/> oseba	158	43.35
<input type="checkbox"/> škarje	21	34.79
<input type="checkbox"/> ženska	60	28.96
<input type="checkbox"/> Bog	22	27.09
<input type="checkbox"/> opera	16	24.28
<input type="checkbox"/> nit	29	24.19
<input type="checkbox"/> žena	24	23.12
<input type="checkbox"/> žival	26	22.23
<input type="checkbox"/> težava	27	17.8
<input type="checkbox"/> stvar	18	16.94

v dajal	313	3.3
<input type="checkbox"/> oseba	75	42.67



# TBL – examples by GDEX

user: Simon Krek corpus: Fida PLUS 620m (SLD sketch grammar)

[Concordance](#)  
[Word List](#)  
[Word Sketch](#)  
[Thesaurus](#)  
[Sketch-Diff](#)  
[? Help on this menu](#)

[Switch menu position](#)

## Tickbox Lexicography - Select Examples

Lemma: **ljubiti**

Gramrel: **kako-kdaj?**

Template: **fidaplus\_slovene** Alternative GDEX configuration:

### strastno

- Ali pa v hudi moški nuji svojo ljubico strastno **ljubi** kar zadaj, za težko zametno zaveso.
- Skupaj sva kar precej popila, nakar sva se strastno **ljubila** dve uri.
- Suh in dolgolas je bil intelektualec, strastno je **ljubil** športne avtomobile, a ni imel voznškega izpita.

### neizmerno

- Bil je pravi Italijan in je poleg dobrih avtov in cigar neizmerno **ljubil** ženske.
- Bil je zares iskren in dober prijatelj, ki je neizmerno **ljubil** in varoval naravo.
- Kako bi mogel drugače, saj je Evo neizmerno **ljubil**.

### neskončno

- Morje zares neskončno **ljubim**, tam sem » bebavo blažena «.
- Posnemati Boga, ki neskončno **ljubi** končne stvari kot končne.
- Solarna oseba, veselega značaja, ki neskončno **ljubi** življenje in barve!

# TBL – Entry Editor

The screenshot displays the 'Entry Editor' window for the Slovenian dictionary 'SLOLEKS\_A'. The main window is titled 'ČLOVEK LJUDI DRUGEGA ČLOVEKA, BITJE, KRAJ'. The entry for 'ljubiti' is shown in three structural parts:

- a) Struktura:**
  - ▶ kdo ljubi koga
  - ▶ kdo ljubi
  - ▶ kdo ljubi kaj
  - [bog, moški, dekle, ženska, človek] ljubiti
  - ljubiti [Boga, domovino, ženo, žensko, dekle, sovražnika, moškega]
  - Bila je ljubezen, ki sta jo poznala mama in oče in ki jo je Charis našla ob Harryju. Bila je ljubezen, o kakršni je vedno sanjala. 'Ljubim ga!' je hotela zaklicati, a besede niso bile potrebne.
  - Partnerja moramo ljubiti takega, kot je, ne pa da od njega zahtevamo, naj se odreče nečemu, kar je zanj bistveno.
  - Spoznanje, da je sposobna ljubiti, se je počasi priplazilo k njej med dolgimi urami zadnje noči. To je njeno žalost še poglobilo. Nekoga je ljubila in zdaj je mrtev.
  - Ko ljubimo, je razlika v letih povsem nepomembna.
  - Nenehno flirtamo z ničem, smrt je naša sestra, lahko pa ljubimo in gotovo se naša ljubezen skriva nekje v tem kaosu.
  - Čeprav svobode ne potrebujejo nič manj kot njihovi partnerji, so nagnjeni k ljubosumju. So tudi takšni, ki ljubijo na izrazito strelčevski način.
  - Moč, ki ti daje poguma, da v življenju poskušaš vedno misliti na to, da te Bog ljubi.
- b) Struktura: rbz GBZ**
  - [brezpogojno, zares, strastno, neizmerno, neskončno, nadvse, iskreno, srčno] ljubiti
  - Najlepša lastnost človeka je, da je sposoben brezpogojno ljubiti
  - " Ne vem, ali je v življenju mogoče zares ljubiti dve ženski; dveh domovin prav gotovo ne moreš ljubiti.
  - Skupaj sva kar precej popila, nakar sva se strastno ljubila dve uri.
  - Bil je zares iskren in dober prijatelj, ki je neizmerno ljubil in varoval naravo.
  - Solarna oseba, veselega značaja, ki neskončno ljubi življenje in barve!
- c) Struktura: gbz GBZ**
  - [znati, nehati, naučiti se] ljubiti
  - Kaj pa je tisto, kar najbolj ceni pri svoji Cvetki? » Najbolj mi je všeč, da me podpira pri mojem delu. Cenim njeno odločnost, doslednost. Je oseba, ki zna ljubiti in biti ljubljena.
  - Najbrž ga ni nehala ljubiti, četudi se je poročila z drugim moškim.
  - Tako niso imeli možnosti, da bi spoznali ljubezen, da bi se naučili ljubiti.
  - Zato se človek, ki zna ljubiti, ne boji, da bi v zakonu izgubil svobodo.

**skladenjske zveze:**  
ljubiti z (vsem) srcem

- Ko sem ga ljubila, sem ljubila z vsem srcem, z vso dušo.
- Ljubi s srcem in ne pričakuj, da ti bo ljubezen vrnila.

**ljubiti koga do konca (življenja)**

The right pane shows the morphological analysis of the word 'ljubiti' with various grammatical forms and their corresponding morphological tags (e.g., ZGLED, KOLOKACIJA, ZGLEDI).

# TBL - Animation

**Entry Editor : Leksikalna baza slovensčine A [SLOLEKS\_A] (DPS SS.)**

User: simon.krek

Project: SLOLEKS\_A  
Headword: ljubiti  
Senseid: 4  
Words: 36  
Characters: 187

NAVIGATION PANEL  
Quick Edit  
Upload Select all Lose edits

ljubiti  
1 o čustveni navezanosti  
1.1 o vsadnem ljubavi  
1.2 o sprejemanju ali občudovanju  
2 ?o ljubezni do stvari, pojavov, dogodkov/ o življenju v dem  
3 o pripravljenosti za dejanje  
4 spolno občevati

SEARCH  
ljubiti  
Download Select all

ljubiti ES

**ČLOVEK LJUDI DRUGEGA ČLOVEKA, BITJE, KRAJ**

a) Struktura:

- ▶ kdo ljubi koga
- ▶ kdo ljubi
- ▶ kdo ljubi kaj

■ [bog, moški, dekle, ženska, človek] ljubiti

■ ljubiti [Boga, domovino, ženo, žensko, dekle, sovražnika, moškega]

- Bila je ljubezen, ki sta jo poznala mama in oče in ki jo je Charis našla ob Harryju. Bila je ljubezen, o kakršni je vedno sanjala. 'Ljubim ga!' je hotela zaklicati, a besede niso bile potrebne.
- Partnerja moramo ljubiti takega, kot je, ne pa da od njega zahtevamo, naj se odreče nečemu, kar je zanj bistveno.
- Spoznanje, da je sposobna ljubiti, se je počasi priplazilo k njej med dolgimi urami zadnje noči. To je njeno žalost še poglobilo. Nekoga je ljubila in zdaj je mrtev.
- Ko ljubimo, je razlika v letih povsem nepomembna.
- Nenehno flirtamo z ničem, smrt je naša sestra, lahko pa ljubimo in gotovo se naša ljubezen skriva nekje v tem kaosu.
- Čeprav svobode ne potrebujejo nič manj kot njihovi partnerji, so nagnjeni k ljubosumju. So tudi takšni, ki ljubijo na izrazito strelčevski način.
- Moč, ki ti daje poguma, da v življenju poskušaš vedno misliti na to, da te Bog ljubi.

b) Struktura: rbz GBZ

■ [brezpogojno, zares, strastno, neizmerno, neskončno, nadvse, iskreno, srčno] ljubiti

- Najlepša lastnost človeka je, da je sposoben brezpogojno ljubiti
- " Ne vem, ali je v življenju mogoče zares ljubiti dve ženski; dveh domovin prav gotovo ne moreš ljubiti.
- Skupaj sva kar precej popila, nakar sva se strastno ljubila dve uri.
- Bil je zares iskren in dober prijatelj, ki je neizmerno ljubil in varoval naravo.
- Solarna oseba, veselega značaja, ki neskončno ljubi življenje in barve!

c) Struktura: gbz GBZ

■ [znati, nehati, naučiti se] ljubiti

- Kaj pa je tisto, kar najbolj ceni pri svoji Cvetki? » Najbolj mi je všeč, da me podpira pri mojem delu. Cenim njeno odločnost, doslednost. Je oseba, ki zna ljubiti in biti ljubljena.
- Najbrž ga ni nehala ljubiti, četudi se je poročila z drugim moškim.
- Tako niso imeli možnosti, da bi spoznali ljubezen, da bi se naučili ljubiti.
- Zato se človek, ki zna ljubiti, ne boji, da bi v zakonu izgubil svobodo.

**skladenjske zveze:**  
ljubiti z (vsem) srcem

- Ko sem ga ljubila, sem ljubila z vsem srcem, z vso dušo.
- Ljubi s srcem in ne pričakuj, da ti bo ljubezen vrnila.

ljubiti koga do konca (življenja)

<clanek><geslo><pomen><skladenjska\_skupine><zaledi><zaledi>

ZGLED Partnerja moramo ljubiti takega, kot je, ne pa da od njega zahtevamo, naj se odreče nečemu, kar je zanj bistveno.

ZGLED Spoznanje, da je sposobna ljubiti, se je počasi priplazilo k njej med dolgimi urami zadnje noči. To je njeno žalost še poglobilo. Nekoga je ljubila in zdaj je mrtev.

ZGLED Ko ljubimo, je razlika v letih povsem nepomembna.

ZGLED Nenehno flirtamo z ničem, smrt je naša sestra, lahko pa ljubimo in gotovo se naša ljubezen skriva nekje v tem kaosu.

ZGLED Čeprav svobode ne potrebujejo nič manj kot njihovi partnerji, so nagnjeni k ljubosumju. So tudi takšni, ki ljubijo na izrazito strelčevski način.

ZGLED Moč, ki ti daje poguma, da v življenju poskušaš vedno misliti na to, da te Bog ljubi.

SKLADENJSKA\_STRUKTURA

STRUKTURA rbz GBZ

KOLOKACIJE

KOLOKACIJA K brezpogojno, zares, strastno, neizmerno, neskončno, nadvse, iskreno, srčno /K ljubiti

ZGLEDI

ZGLED Najlepša lastnost človeka je, da je sposoben brezpogojno ljubiti

ZGLED " Ne vem, ali je v življenju mogoče zares ljubiti dve ženski; dveh domovin prav gotovo ne moreš ljubiti.

ZGLED Skupaj sva kar precej popila, nakar sva se strastno ljubila dve uri.

ZGLED Bil je zares iskren in dober prijatelj, ki je neizmerno ljubil in varoval naravo.

ZGLED Solarna oseba, veselega značaja, ki neskončno ljubi življenje in barve!

SKLADENJSKA\_STRUKTURA

STRUKTURA gbz GBZ

KOLOKACIJE

KOLOKACIJA K znati, nehati, naučiti se /K ljubiti

ZGLEDI

ZGLED Kaj pa je tisto, kar najbolj ceni pri svoji Cvetki? » Najbolj mi je všeč, da me podpira pri mojem delu. Cenim njeno odločnost, doslednost. Je oseba, ki zna ljubiti in biti ljubljena.

ZGLED Najbrž ga ni nehala ljubiti, četudi se je poročila z drugim moškim.

ZGLED Tako niso imeli možnosti, da bi spoznali ljubezen, da bi se naučili ljubiti.

ZGLED Zato se človek, ki zna ljubiti, ne boji, da bi v zakonu izgubil svobodo.

# GDEX – Good Dictionary Examples

- system for evaluation (ranking) of sentences with respect to their suitability to serve as dictionary examples
- sorting sentences so that good examples do not have to be searched for in hundreds of unusable sentences
- initially trained on English, but it did not give good results for other languages

# GDEX – configuration

- parameters in a GDEX configuration file
- GDEX Tools web-interface to create and use custom GDEX configurations
- the GDEX evaluation process
  - ranking of out-of-corpus sentences
  - evaluation of TBLex logs
  - cooperation with WEKA

# GDEX classifiers

- procedures that quantify measurable features of sentences or tokens
- sentence classifiers: sentence length, keyword position, etc.
- token classifiers: token frequencies, matches to RE, etc.

# Evaluation of TBLex logs

user: Simon Krek corpus: Fida PLUS 620m (SLD sketch grammar)

[Concordance](#)  
[Word List](#)  
[Word Sketch](#)  
[Thesaurus](#)  
[Sketch-Diff](#)  
[? Help on main menu](#)

[Switch menu position](#)

## Textbox Lexicography - Select Examples

Lemma: **ljubiti**

Gramrel: **kako-kdaj?**

Template: **fidaplus\_slovene** Alternative GDEX configuration:

### GDEX: Slovene3

### GDEX: default configuration

#### strastno

- Ali pa v hudi moški nuji svojo ljubico strastno **ljubi** kar zadaj, za težko žametno zaveso.
- Skupaj sva kar precej popila, nakar sva se strastno **ljubila** dve uri.
- Suh in dolgolas je bil intelektualec, strastno je **ljubil** športne avtomobile, a ni imel voznškega izpita.

#### strastno

- Sophie me zdaj strastno **ljubi** .
- Nate, ki te strastno **ljubim** .
- Vse dneve je pridno delal in jo vse noči strastno **ljubil** .

#### neizmerno

- Bil je pravi Italijan in je poleg dobrih avtov in cigar neizmerno **ljubil** ženske.
- Bil je zares iskren in dober prijatelj, ki je neizmerno **ljubil** in varoval naravo.
- Kako bi mogel drugače, saj je Evo neizmerno **ljubil** .

#### neizmerno

- Dekle, ki neizmerno **ljubi** .
- Sam je Marijo neizmerno **ljubil** .
- Bog nas neizmerno **ljubi** .

#### neskončno

- Morje zares neskončno **ljubim** , tam sem » bebavo blažena «.
- Posnemati Boga, ki neskončno **ljubi** končne stvari kot končne.
- Solarna oseba, veselega značaja, ki neskončno **ljubi** življenje in barve!

#### neskončno

- Tudi ona ga neskončno **ljubi** .
- Neskončno te **ljubim** .
- Zavedel se je, kako neskončno jo **ljubi** .

# Cooperation with WEKA

The screenshot displays the Weka Explorer application window. The interface includes a menu bar with options: Preprocess, Classify, Cluster, Associate, Select attributes, and Visualize. Below the menu bar are buttons for Open file..., Open URL..., Open DB..., Generate..., Undo, Edit..., and Save....

The main area is divided into several sections:

- Filter:** A dropdown menu set to "None" with an "Apply" button.
- Current relation:** Shows "Relation: Logged\_Sentences\_from\_TBLEx" and "Instances: 47494".
- Attributes:** A list of attributes with checkboxes: average (checked), max, min, sentence length, sentence length score, and good. Buttons for "All", "None", "Invert", and "Pattern" are also present.
- Selected attribute:** Displays statistics for the "average" attribute: Name: average, Type: Numeric, Missing: 0 (0%), Distinct: 43691, Unique: 40610 (86%). A table below shows the following statistics:

Statistic	Value
Minimum	7
Maximum	23512219.5
Mean	6655208.193
StdDev	2537532.627
- Visualize:** A histogram showing the distribution of the "average" attribute. The x-axis ranges from 7 to 23512219.5, with a mean value of 11756113.25. The histogram is colored blue and red.

The status bar at the bottom shows "Status OK" and a "Log" button.



# New GDEX configuration

user: Simon Krek used tokens: 586,435 / 1,000,000 days left: 199

## Corpora

[+ Create corpus](#)

[+ WebBootCaT](#)

[Configuration templates](#)

[Sketch grammars](#)

[User groups](#)

## Admin

[Local administration](#)

## Support

[Help](#)

[Report a bug](#)











## GDEX configurations

### My GDEX configurations

Filename	Name
no GDEX configurations	

[+ Add new GDEX configuration](#)

### Other users' GDEX configurations

Filename	Name	Owner		
Slovene1_beta_compatible_2	Slovene1	Milos Husak		
Slovenian01b_beta_compatible_2	Slovene1b	Milos Husak		
Slovenian02b_beta_compatible_2	Slovene2	Milos Husak		
Slovenian03_beta_compatible_2	Slovene3	Milos Husak		
Slovenian03b_beta_compatible_3	Slovene3b	Milos Husak		

# Tomorrow

- Automatic extraction
  - Sketch Grammar
  - GDEX
- Proposal for the compilation of the new dictionary of modern Slovene