

User manual

MORF-HR application

Morphological Lexicon of the Croatian Language

Contents

1. Introduction.....	4
2. General notes.....	5
2.1. Interface language selection.....	5
2.2. Lexicon statistics.....	5
2.3. Cookies	5
3. Search.....	6
3.1. Opening the search page.....	6
3.2. Basic search	6
3.3. Advanced filters.....	6
3.3.1. Part of speech.....	6
3.3.2. Verb type and class.....	6
3.3.3. Search by inflectional forms	7
3.3.4. Search by root only.....	7
3.3.5. Root variance.....	7
3.3.6. Prefix and suffix (Derivational Lexicon)	7
3.4. Active filters.....	7
3.5. Clearing filters.....	7
3.6. Search results	7
3.7. Pagination.....	8
4. Detailed lemma view.....	9
4.1. Basic information.....	9
4.2. Inflectional forms (tab F).....	9
4.2.1. Nouns.....	9
4.2.2. Adjectives	9
4.2.3. Verbs.....	9
4.2.4. Other parts of speech.....	9
4.3. Derivational data (tab T).....	9
4.4. Derivational family (modal).....	10
4.5. Back to results	10
5. Login and logout (authorized users).....	11
5.1. Login	11
5.2. Logout.....	11
6. Administrator functions.....	12
6.1. Importing data from Excel files.....	12
6.1.1. File selection.....	12
6.1.2. Starting the import.....	12

6.1.3. Import results	12
6.2. Deleting lemmas.....	12
6.3. Refreshing statistics.....	12
6.4. Identifier column in the results table	13
7. Frequently Asked Questions and Notes	14
7.1. Why are there always capital letters in the search field?.....	14
7.2. What does the asterisk (*) next to a lemma mean?.....	14
7.3. Difference between the Inflectional (F) and Derivational Lexicon (T).....	14
7.4. What is a derivational family?	14
7.5. Search yields no results.....	14

1. Introduction

MORF-HR is a linguistic resource that comprises the Derivational and Inflectional Lexicon of the Croatian language. It enables searching for lemmas, viewing inflectional forms (declension, conjugation, comparison), and displays the word-formation relations between words as well as their morphological structure.

MORF-HR is available to all users without registration.

2. General notes

2.1. Interface language selection

The interface language is changed by clicking on the drop-down menu with a flag in the upper right corner of the navigation bar. Two languages are available:

- Croatian flag (HR) — Croatian interface
- British flag (EN) — English interface

The selection is applied immediately and remains active for the entire session.

2.2. Lexicon statistics

By clicking on the **(i)** icon in the navigation bar, a window with lexicon statistics opens. The following is displayed:

- Total number of lemmas in the Inflectional Lexicon
- Distribution of lemmas by part of speech in the Inflectional Lexicon
- Total number of lemmas in the Derivational Lexicon
- Distribution of lemmas by part of speech in the Derivational Lexicon

2.3. Cookies

Upon the first visit, a cookie notification is displayed at the bottom of the screen. Clicking the "OK" button permanently closes the notification (for 365 days).

3. Search

3.1. Opening the search page

The application opens directly on the search page. In the navigation bar, clicking on the name MORF-HR always returns the user to the search page.

3.2. Basic search

In the search field (the central field at the top of the page), the term to be searched is entered.

Input rules:

- The input is automatically converted to uppercase
- Minimum length: 1 character; maximum length: 30 characters
- To enter the special character Ě (yat), use the Ě button to the right of the field
- Up to 3 characters: searches for an exact match; 4 or more characters: searches for a partial match (the term can be located anywhere within the lemma)

The search is initiated by clicking the "Search" button or by pressing the *Enter* key.

3.3. Advanced filters

By clicking "Advanced filters" below the search field, a panel with additional criteria for narrowing down the results is expanded.

3.3.1. Part of speech

A drop-down menu for selecting the part of speech. Available categories:

- Noun
- Adjective
- Pronominal adjective
- Verb
- Pronoun (personal, reflexive, demonstrative, possessive, interrogative, interrogative/relative)
- Numeral (cardinal, ordinal)
- Adverb
- Past participle
- L-participle
- Present verbal adverb
- Past verbal adverb
- Preposition
- Numeral noun
- Interjection
- Particle
- Conjunction

Note: Selecting a part of speech is not mandatory, but it narrows the set of results and speeds up the search.

3.3.2. Verb type and class

Available only for the part of speech *Verb*. After selecting *Verb* in the "Part of speech" filter, a dropdown menu with verb types and classes becomes active:

I.1., I.2., I.3., I.4., I.5., I.6., I.7., II., III.1., III.2., IV., V.1., V.2., V.3., V.4., VI., VII.

3.3.3. Search by inflectional forms

The "Search by inflectional forms" checkbox includes inflectional forms (case endings, verb tenses, etc.) in the search, rather than just the base lemma forms. This option is mutually exclusive with the "Search by root only" option.

3.3.4. Search by root only

The "Search by root only" checkbox restricts the search to root morphemes. Activating this option disables "Search paradigms" and activates the "Root variant" dropdown menu.

3.3.5. Root variance

Available only when "Search by root only" is enabled. A dropdown menu with values 1–5 for filtering by a specific root variant. The root variant refers to homographic roots, e.g., KOS_1, KOS_2, KOS_3, etc., where KOS_1 includes words morphologically related to the adjective *kos* (slanted, sloped), KOS_2 words related to the noun *kos* (blackbird), KOS_3 to the noun *kosa* (hair on the head) ...

3.3.6. Prefix and suffix (Derivational Lexicon)

Search fields by morphemes in the Derivational Lexicon:

- Prefix: prefix entry (minimum 1 character)
- Suffix: suffix entry (minimum 1 character); the special character $\emptyset$ ($\emptyset$ button to the right of the field) is used for lemmas without a suffix

The database can be searched by a single prefix, a single suffix, or by a combination of one prefix and one suffix. Searching by a prefix and/or suffix refers to their base, non-alternated forms—that is, their form as written in the deep structure. The search results will yield all allomorphs of a given affix. This means that a search for the prefix *-iz*, for example, will return lemmas such as *izgraditi*, *ispisati*, *izabrati*, *iščupati*, but also *reizabrati*.

If the user wants to search for combinations of affixes, e.g., *-teljstvo*, this combination should be entered in the main search field (the central field at the top of the page — *Search term*).

Note: Filling in the prefix or suffix field is a sufficient criterion to initiate a search — entering a term in the main field is not required.

3.4. Active filters

After a search is initiated, active filters are displayed above the results as colored pills. Each pill shows the filter name and the selected value. Clicking the × within a pill removes that filter, and the search is automatically repeated with the remaining filters.

By clicking the "Clear all" button, all filters are removed and the view returns to the initial state.

3.5. Clearing filters

The "Clear filters" button (below the search field, within the form) clears all entered filters and resets the form to its initial state...

3.6. Search results

The results are displayed in a table with the following columns:

- **Lemma** — clickable lemma name; if a superscript is displayed next to the lemma (e.g., kuća²), it is an identifier used to distinguish homographic lemmas
- **Root** — the lexical morpheme of the lemma; the superscript indicates the root variant
- **Part of speech** — Part of speech
- **Type** — type/class (e.g., verb type); clickable for verbs — clicking sets the Type/Class filter to the selected Type/Class
- **Lexicon** — indicators of presence in the lexicons:
 - F (gray indicator) — the lemma is in the Inflectional Lexicon
 - T (blue indicator) — the lemma is in the Derivational Lexicon

Non-existent lemmas are marked with an asterisk () — these are words that do not exist in (contemporary) standard Croatian but are used in the word-formation model.*

By clicking on the lemma name, the detailed view is opened.

3.7. Pagination

Below the table, controls are displayed containing:

- Information regarding the number of displayed and total results, as well as the current page
- Page navigation (« previous, page numbering, next »)
- Each page displays up to 20 results

4. Detailed lemma view

4.1. Basic information

At the top of the page, the following is displayed:

- Lemma (with an identifier as a superscript, if available)
- Root
- Part of speech
- Type/class (if multiple classes are available))

4.2. Inflectional forms (tab F)

Displays all inflectional forms of the lemma.

4.2.1. Nouns

The table displays the following columns: Gender, Number, Case, Class, Paradigm.

4.2.2. Adjectives

The table displays the following columns: Gender, Number, Degree (definite/indefinite), Case, Type, Paradigm.

Below the main table, a separate table with comparative and superlative forms is displayed (if available).

4.2.3. Verbs

The table displays the following columns: Form, Number, Person, Class, Paradigm.

Verb tenses and forms: *Present, Future I, Future II, Aorist, Imperfect, Perfect, Pluperfect, Imperative, Conditional I, Conditional II*.

Below the forms table, there is a section with verb participles.

4.2.4. Other parts of speech

A generic table with the following columns: Gender, Number, Case, Type, Paradigm.

4.3. Derivational data (tab T)

The T tab displays word-formation data for the lemma (available only if the lemma exists in the derivational database). A derivational card is displayed, containing:

- Word-formation chain — a graphical representation (SVG diagram) of the word-formation sequence from the base lemma to the selected lemma
- Lemma and part of speech
- Morphological structure (surface level) — a color-coded graphical representation: roots (orange), prefixes (green), suffixes (blue)
- Morphological structure (deep level) — a color-coded graphical representation: roots (orange), prefixes (green), suffixes (blue)
- Word-formation pattern — a formula showing the word-formation base and affixes
- Word-formation process — specification of the process (e.g. suffixation, prefixation)
- "Derivational family" button — opens a modal with a diagram of the entire derivational family

4.4. Derivational family (modal)

By clicking the "Word-formation family" button, a modal window opens with a graphical (SVG) representation of the entire derivational family tree to which the lemma belongs. The diagram is movable, and it can be downloaded in PNG or JPG format.

4.5. Back to results

The "Back" button (or the browser's Back button) returns to the previous search results view, preserving all applied filters and the entered search term.

5. Login and logout (authorized users)

5.1. Login

By clicking the person icon and the "Login" text in the upper right corner of the navigation bar, a login modal opens.

Enter:

- Username
- Password
- "Remember me" option (keeps you logged in for 14 days)

Clicking the "Login" button triggers the system to verify the credentials. In case of incorrect data, an error message is displayed within the modal.

After a successful login, the browser is redirected to the data import page.

5.2. Logout

Clicking the "Logout" button (right corner of the navigation bar, visible only to logged-in users) logs the user out of the system.

6. Administrator functions

The following functions are available exclusively to users with administrator privileges.

6.1. Importing data from Excel files

The import page is accessible by clicking "Import" in the navigation bar (visible only to administrators).

6.1.1. File selection

Clicking the "Select files" button opens a dialog box for selecting one or more Excel files (.xlsx or .xls). Up to 20 files can be selected simultaneously (or a different limit if configured). The status of the selected files is displayed below the button.

6.1.2. Starting the import

Clicking the "Import" button starts the import process. An animated spinner is displayed during processing.

The application processes each file individually and displays a table with the results.

6.1.3. Import results

The following is displayed for each file:

- File name
- Processing status:
 - Imported (green color) — the file was successfully processed
 - Empty file (red color) — the file contains no content
 - Invalid format (red color) — the file is not in the correct format
 - Failed recognition (red color) — unable to determine the part of speech
 - Processing error (red color) — an unexpected error occurred
- Messages — error details (or a dash if there are no messages)

6.2. Deleting lemmas

In the search results table, each row contains at least one delete button (a trash can icon, a tree view icon, or both icons). The trash can icon triggers the deletion of a specific lemma from the Inflectional Lexicon, while clicking the tree view icon deletes **all lemmas** from the Derivational Lexicon that belong to the same root as the lemma in that row.

Procedure:

1. Locate the lemma to be deleted using the search function
2. In the lemma's row, click the button with the trash can icon or the tree diagram icon
3. Confirm the deletion in the confirmation dialog box
4. After a successful deletion, a green notification appears: *Lemma successfully deleted*
5. In case of an error, a red notification appears with a description of the error

6.3. Refreshing statistics

In the Statistics window (opened by clicking the (i) icon in the navigation bar), a "Refresh statistics" button is available (visible only to administrators).

Clicking the button recalculates the statistics. In case of success, a green text is displayed, and in case of an error, a red text with a description of the error appears.

6.4. Identifier column in the results table

An additional column with the values of internal root and lemma identifiers is visible to administrators in the search results table. This column is not visible to unauthorized users. The value of the identifier is used to update data for a specific lemma by entering that identifier into the "INDIKATOR" column in the Excel file next to the lemma data. An Excel file filled out this way is processed in a manner that updates the lemma data in the database that matches the identifier value.

7. Frequently Asked Questions and Notes

7.1. Why are there always capital letters in the search field?

Lemmas in the database are stored in capital letters (the standard for morphological lexicons). The system automatically converts the input into capital letters for consistent searching.

7.2. What does the asterisk (*) next to a lemma mean?

An asterisk indicates a virtual lemma — a form that does not exist in the contemporary standard Croatian language but is used as a stem in word-formation descriptions.

7.3. Difference between the Inflectional (F) and Derivational Lexicon (T)

- The Inflectional Lexicon (F) contains all the inflectional forms of a lemma: noun, adjective, pronoun and verb paradigms, etc.
- The Derivational Lexicon (T) describes the morphological structure of lemmas and the derivational relations between them. It contains roots, stems and derivatives, which affixes were used, which derivational processes were applied, etc.

A lemma can be present in only one or in both lexicons.

7.4. What is a derivational family?

A derivational family is a set of all lemmas that are derivationally related and share the same lexical morpheme (root). The graphical representation (tree) shows the hierarchical organization of derivational families.

7.5. Search yields no results

Please check:

1. Whether a term of at least 1 character has been entered, whether a part of speech has been selected, or whether a prefix/suffix has been entered.
2. Whether the term is spelled correctly (without typos and diacritics if the searched lemma does not have diacritics).
3. Whether the "Type" filter is too restrictive — try removing some filters.