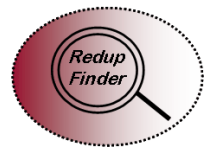


g r a z database on reduplication

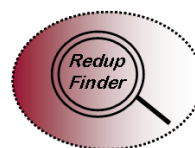


<http://ling.uni-graz.at/reduplication/>



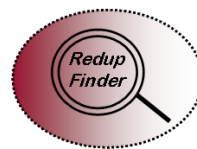
Technical Documentation of Reduplication Finder 1.1

Programmed and written
by
Dipl.-Ing. Angela Fessel
(angela.fessel@uni-graz.at)



1 Table of contents

<u>1</u>	<u>TABLE OF CONTENTS</u>	<u>2</u>
<u>2</u>	<u>INTRODUCTION</u>	<u>3</u>
2.1	GENERAL PURPOSE	3
<u>3</u>	<u>INSTALLATION</u>	<u>5</u>
3.1	SYSTEM REQUIREMENTS	5
3.2	INSTALLATION	5
<u>4</u>	<u>FUNCTIONAL DESCRIPTION</u>	<u>6</u>
4.1	GENERAL	6
4.2	STEPS:	6
4.2.1	STEP 1 – SELECTION OF THE SEARCH TYPE:	6
4.2.2	STEP 2 – FILE SELECTION	7
4.2.3	STEP 3 – START SEARCH:	17
<u>5</u>	<u>INTERNAL DESCRIPTION</u>	<u>21</u>
5.1	WORD LIST FILE	21
5.2	PATTERN FILE	22
5.3	VOWEL FILE	25
5.4	HTML RESULT DIRECTORY	25
5.4.1	RESULT OF THE FREE SEARCH:	26
5.4.2	RESULTS OF THE WORD LIST SEARCH:	29
5.4.3	RESULT OF THE PATTERN SEARCH	30
5.4.4	MODIFICATION OF THE HTML CONTENT	31
<u>6</u>	<u>COPYRIGHT</u>	<u>34</u>
	REDUPLICATION FINDER 1.1	34
	© 2006, DEPARTMENT OF LINGUISTICS GRAZ	34
	SOFTWARE DISCLAIMER AND COPYRIGHT STATEMENT	34



2 Introduction

The Reduplication Finder searches for reduplications in text files. After the search the results will be presented as an HTML page.

This documentation contains a detailed description of how to install the Reduplication Finder, what kinds of searches are available, how it works, and how the results look like.

2.1 General Purpose

The general purpose of the Reduplication Finder is to search either a single file or all the files in a given directory for words containing partial or full reduplication. Only plain text files with the extension “.txt” or HTML files (“.htm” or “.html”) can be searched.

The Reduplication Finder provides three different kinds of searches:

- **Free Search:** All words are found which contain a repetition of a whatever repeated segmental string (i.e. partial or full reduplication). The reduplication can be separated by “-“ or “~”. Double consonants and double vowels will not be found (e.g. “aa” or “kk”).
- **List Search:** This search provides the option of searching for words from a word list provided by the user (for more details see 5.1). All words appearing in this list will be searched for. These words may or may not contain reduplications. This kind of search is case insensitive.

Remark: This search is kept very general so it can be used for all kinds of word searches where a list of words is available; not exclusively for reduplication.

- **Pattern Search:** This search requires a pattern file (for more details see 5.2), which contains search-patterns. All words which match these patterns will be found and presented in the HTML results.

After all three searches an HTML directory with an index.html file is created. (For more details see 0.) The index.html is the starting point for presenting the search results.

Remark: It can only be decided by the user, if the results are indeed “reduplications” or are other arbitrary repetitions (e.g. *banana*).

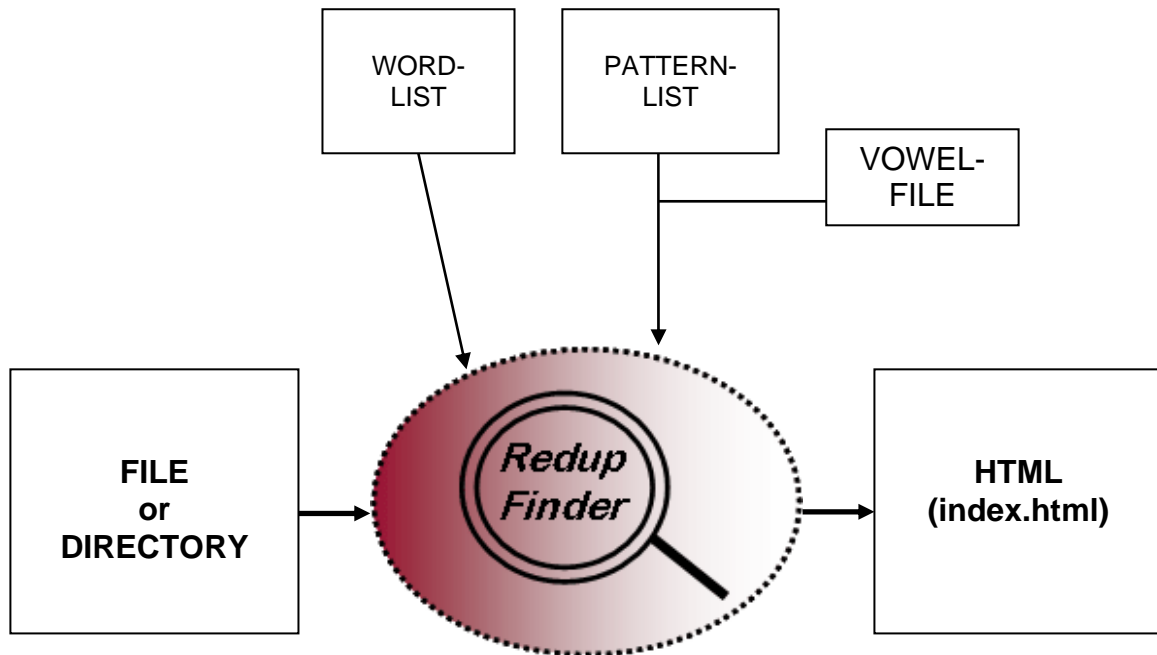
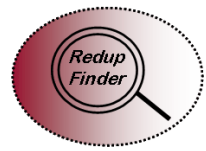
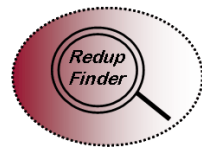


Figure 1: General functionality of Reduplication Finder



3 Installation

3.1 System requirements

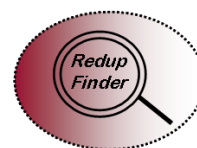
To run the Reduplication Finder 1.1 the following are necessary

- Windows
- Java Runtime Environment 1.5 (JRE)
JRE can be downloaded and installed from: <http://www.java.com/en/download/manual.jsp>

3.2 Installation

1. Save the ReduplicationFinder1.1.zip in a directory
2. Unzip the file ReduplicationFinder1.1.zip
3. Double click on "Reduplication Finder 1.1.exe"

If the program does not start, check if JRE 1.5. is missing.



4 Functional Description

4.1 General

The following pages show in a detailed way how the Reduplication Finder works, how to select one of the three possible searches, which files are necessary, and what the results look like.

4.2 Steps:

4.2.1 Step 1 – Selection of the Search Type:

The first Reduplication Finder window gives the user the option of selecting one of the three kinds of searches:

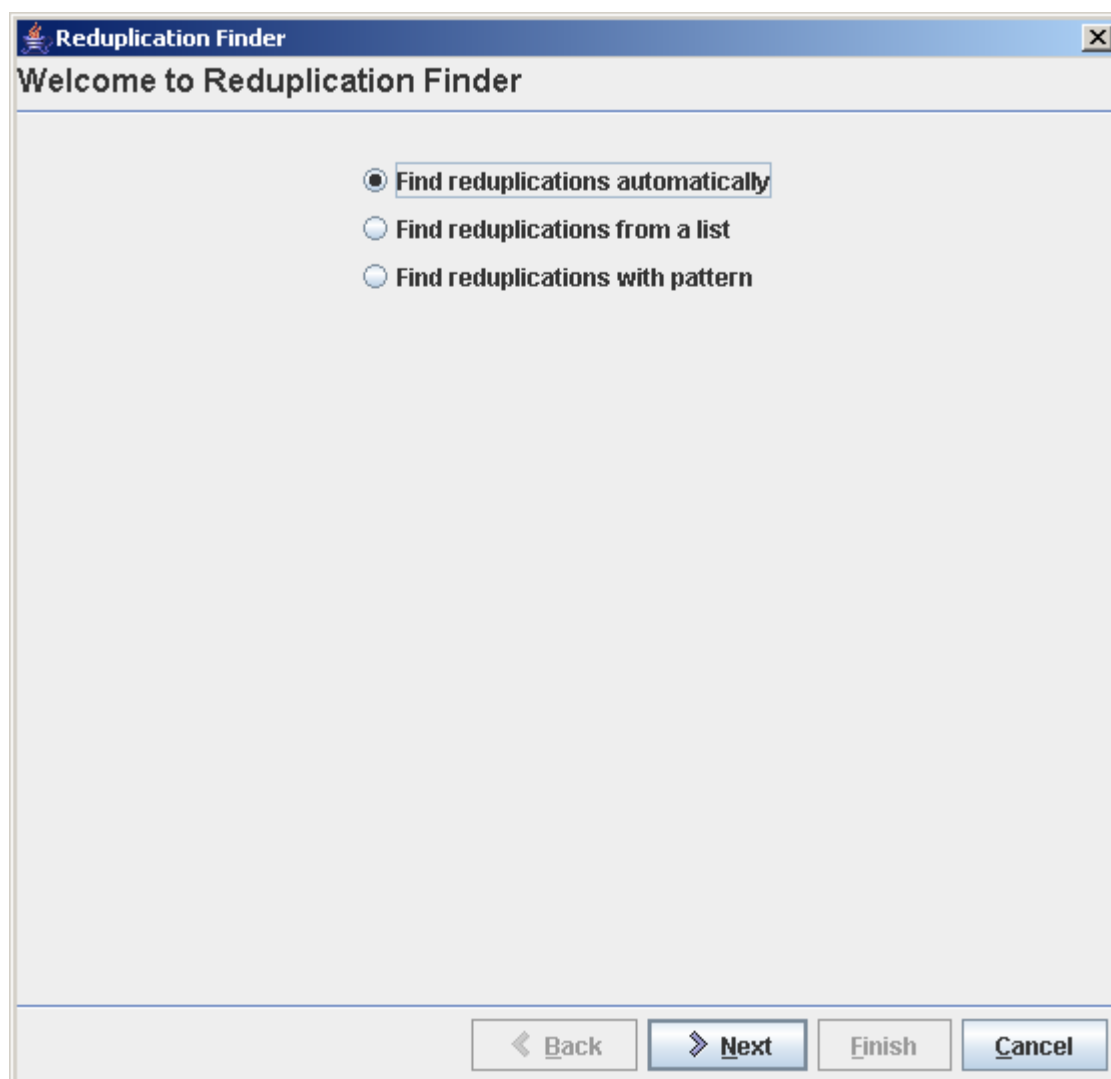
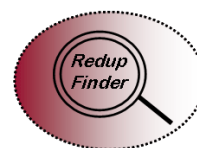


Figure 2: Step 1 of Reduplication Finder



- **Free Search:** Find reduplications automatically.
- **List Search:** Find reduplications from a word list.
- **Pattern Search:** Find reduplications with a pattern list.

4.2.2 Step 2 – File Selection

Depending on the search type selected in the first window, the second window provides for the selection of the files to be searched according to the type of search. For the word list and pattern search the word list or pattern file must also be selected. For the free search only the file or directory to be searched can be selected. For all searches the result directory must be specified.

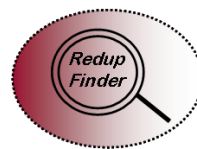
4.2.2.1 Step 2 for a Free Search

For the free search, the second window allows only for the selection of a file or directory and the specification of the result directory:



Figure 3: Step 2 for the free search

For the free search the selection of a file or a directory to be searched and the specification of the



result directory are necessary.

- **Select search file or directory:** Select a file or a directory to be searched. If a directory is chosen, all files and all subdirectories will be searched. To select a file or a directory do the following:
 1. Click either on “Select a file” or “Select a dir”. A file chooser dialog will be opened.

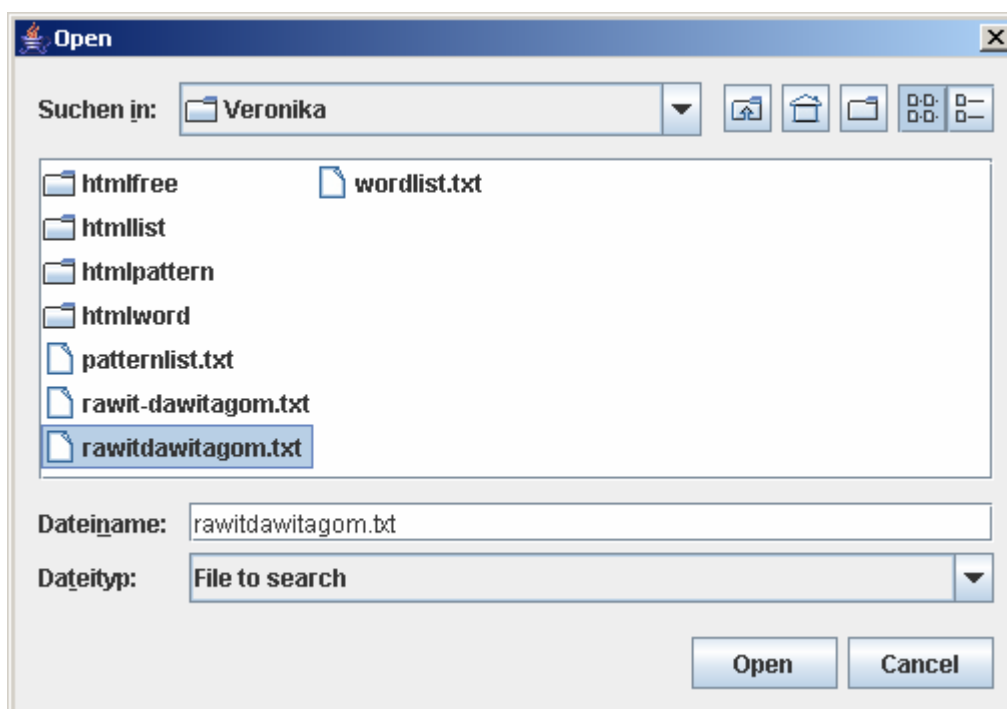


Figure 4: Select a file to search in.

2. Browse to the directory where the file or directory to be searched is available.
3. Press “Return” or click on “Open”.

Remark: It is only possible to select either a file or a directory, not both. If the user tries to select both a file and a directory, the following warning is displayed:

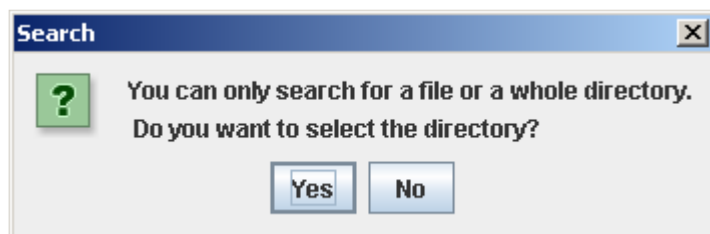
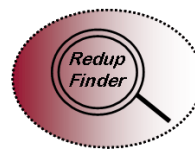


Figure 5: Warning.

If “Yes” is selected, the file will be ignored and the directory will be selected.



- **result directory:** The name of the result directory is automatically set to the directory, where the search file or directory is located and “/html” is added. This is only the default – the user can choose any other directory as to be the result directory.

Remark: If the chosen result file already exists in the result directory, the following warning is displayed:

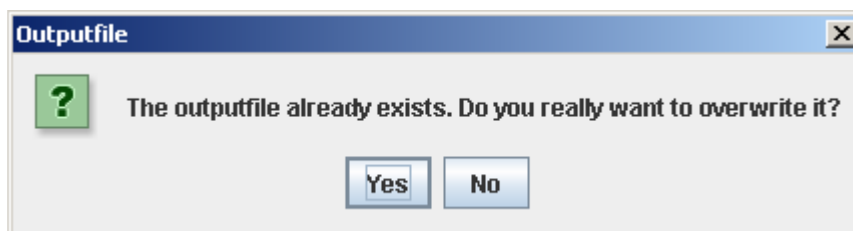


Figure 6: Warning

Remark: After the free search the result directory contains an “index.html” file, which is the starting point of the newly created HTML page. It also contains the following three files “list.html”, “part.html”, “whole.html” and the directories /csjs, /css and /data. For more details regarding the result HTML directory, see 0.

After all fields have been filled in, the Reduplication Finder window should look like this:

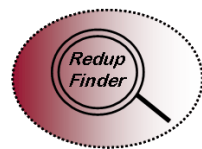
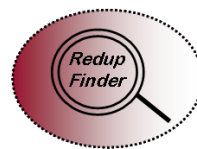


Figure 7: Step 2 after the selection of the files/directories for the free search.



4.2.2.2 Step 2 for a List Search

For the list search the second window looks like this:

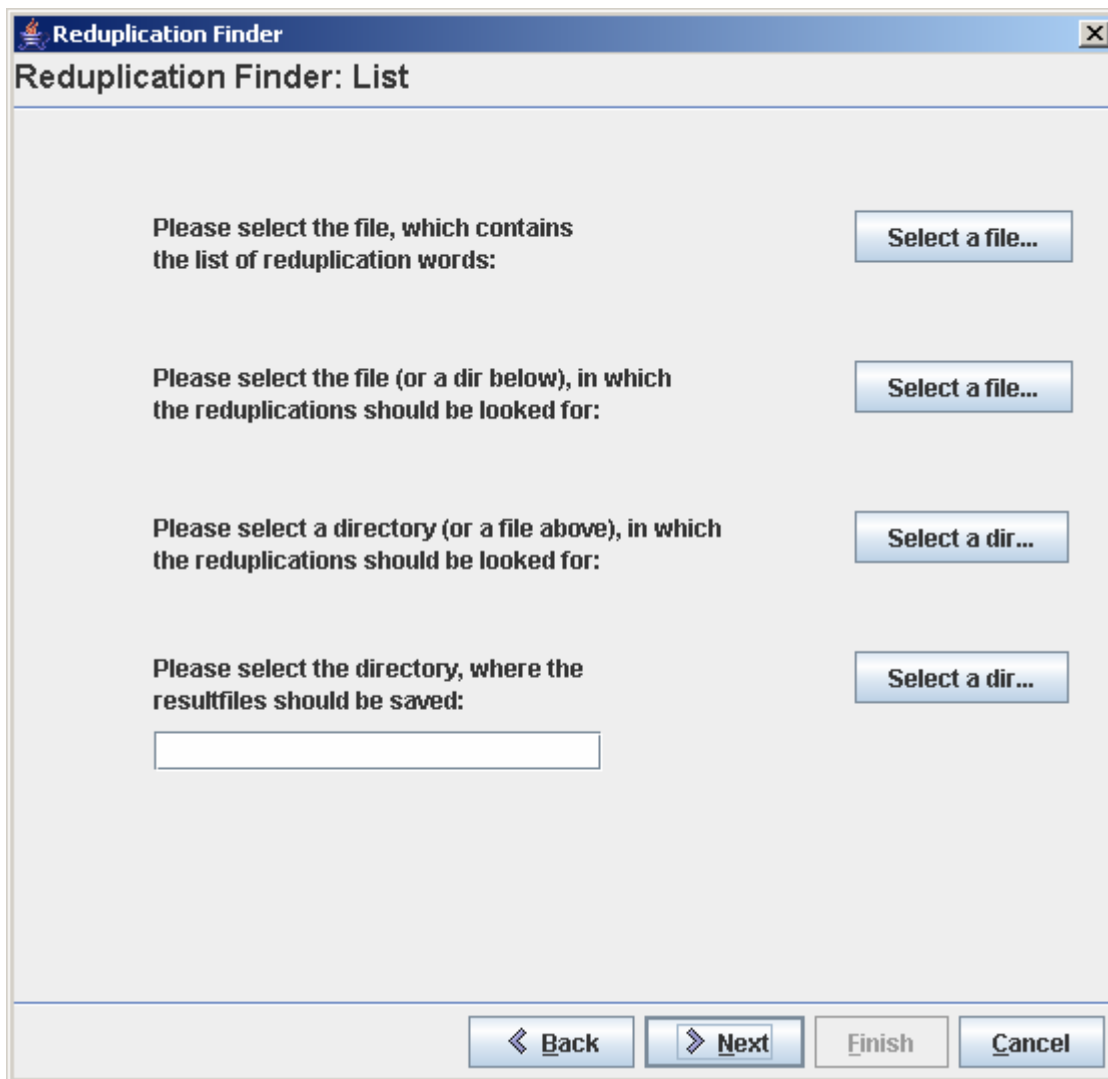


Figure 8: Step 2 of Reduplication Finder

After the selection of the List Search, the word list file, a file or directory to be searched and the output directory are selected here.

- **word list file:** This is a plain text file provided by the user which contains a list of words to be searched for.
To select the word list file do the following:
 1. To choose a ".txt"-file click the "Select a file" button. A file chooser dialog will be opened.

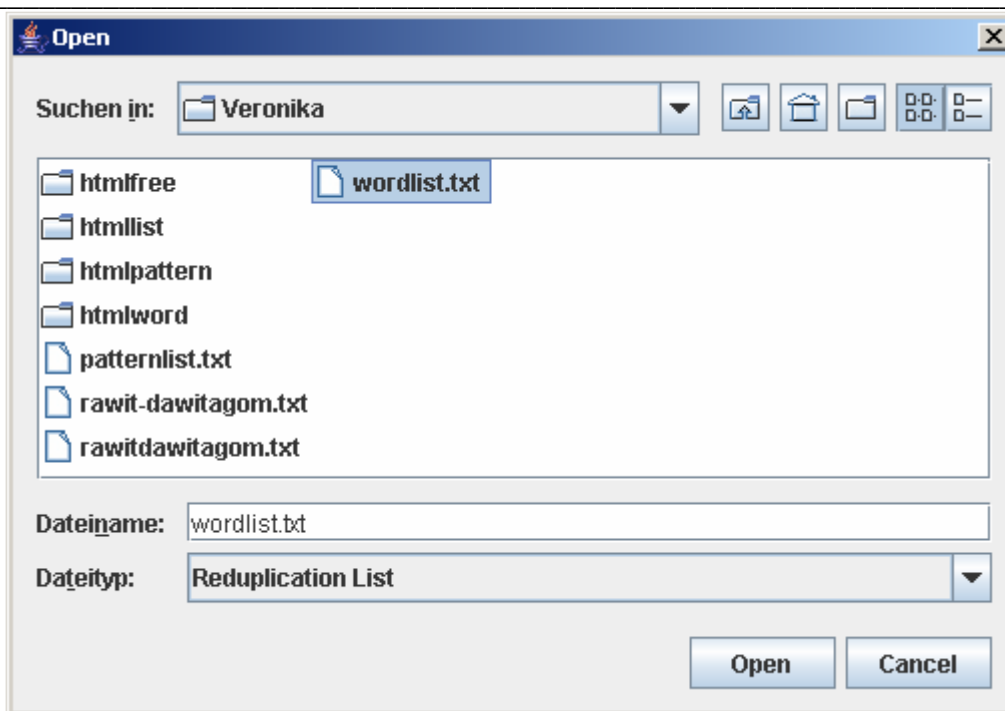
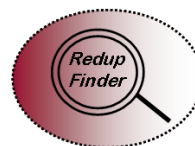


Figure 9: Select a word list file.

2. Browse the directory where the word list file is stored and select it.
3. Press “Return” or click on Open.

Now the word list file is selected.

The word list file consists of words to be searched for. Each line should consist of one word.

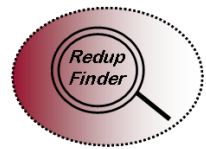
Example for a word list (wordlist.txt):

```
basa-basa
na
parapa
takig-takig
tumog-tumog
tuntom
```

4. **search file or directory:** Select a file or a directory to be searched. If a directory is chosen, all files and all subdirectories will be searched. For instructions on how to select a file or a directory, see the description of Step 2: Free Search above.

Now the file or directory is set correctly.

- **result directory:** The name of the result directory is automatically set to the directory, where the word file was located and “/html” is added. This is only the default – the user can choose any other directory.



After all fields have been filled in, the Reduplication Finder window should look like this:

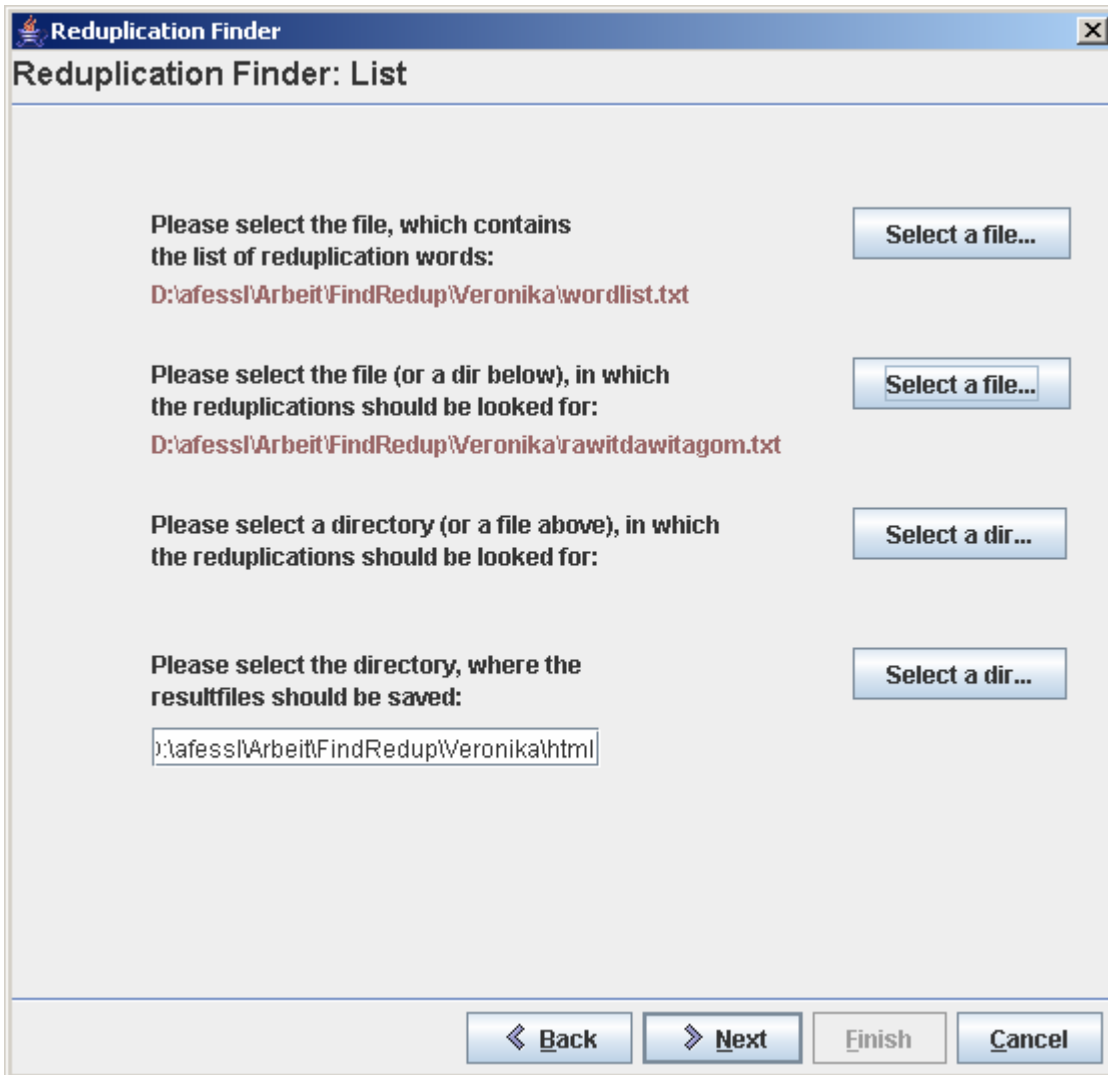
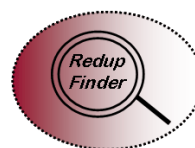


Figure 10: Step 2 after the selection of all files for the list search.



4.2.2.3 Step 2 for a Pattern Search

The second window for the pattern search looks as follows:

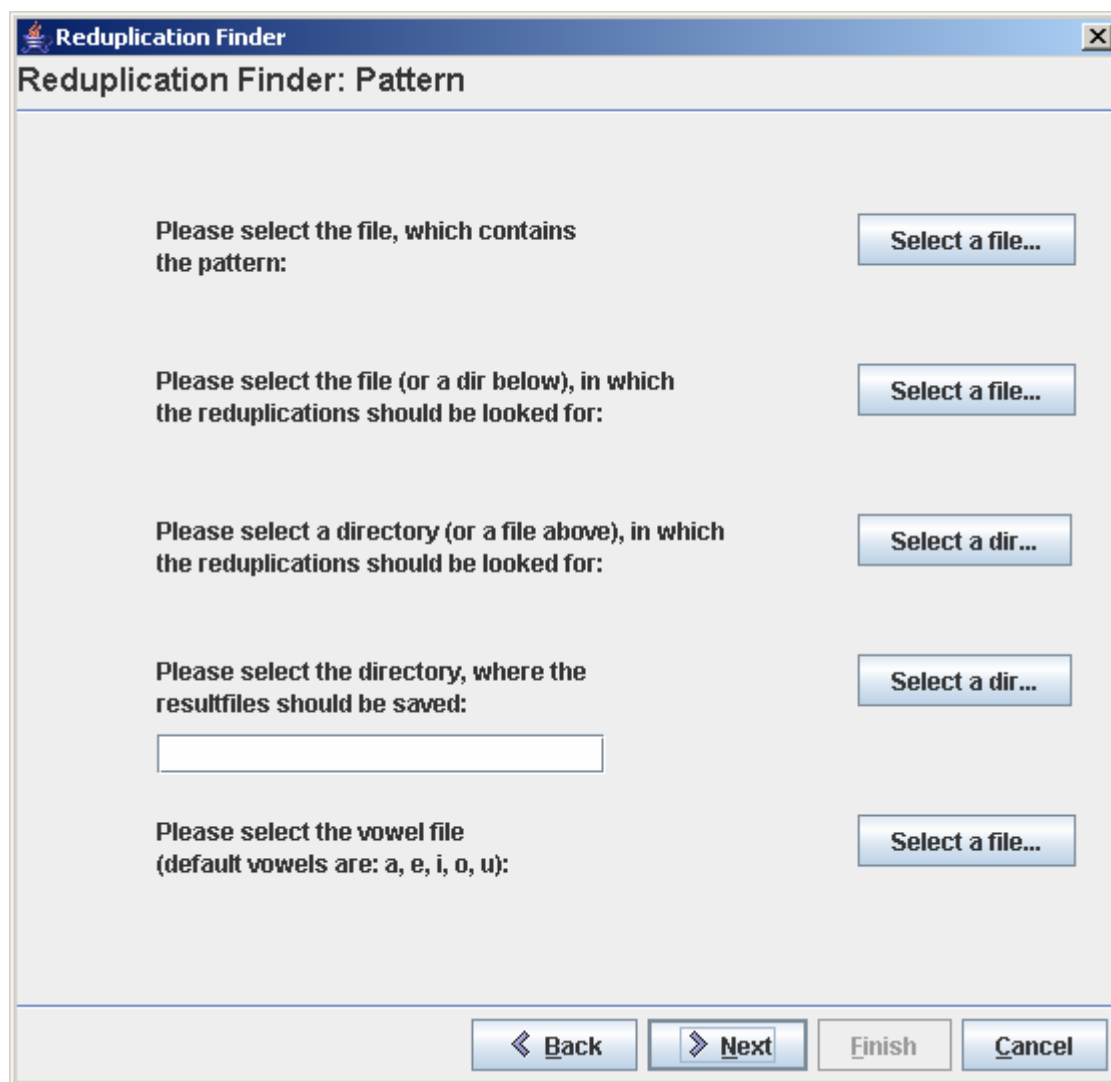


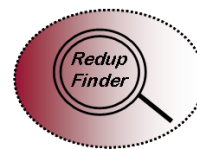
Figure 11: Step 2 for the pattern search.

For the pattern search, the second window allows the user to select a pattern file, a search file or search directory and specify the result directory:

- pattern file:** This is a plain text file provided by the user which contains a list of patterns to be searched for. The selection of the pattern file, can be accomplished the same way as the selection of the word list. This is described in 5.1

The pattern file consists of phonological patterns to be searched for. Each line may consist of only one pattern. For a detailed definition of a pattern file see 5.2.

Example of a pattern file (patternlist.txt):



xV1C1V1x
C1V1C2V1-C1V1C2V1
xV1V1x
xC1V1C2V2C1V1C2V2x

- **search file or directory:** Select a file or directory to be searched through. If a directory is chosen, all files and all subdirectories contained in it will be searched. for instructions on how to select a file or a directory, see the description of Step2: List Search.
- **result directory:** The name of the result directory is automatically set to the directory, where the pattern file is located and "/html" is added. This is only the default – the user can choose any other directory.

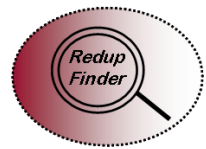
Remark: After the pattern search the result directory contains an "index.html" file, which is the starting point for the newly created HTML page. It also contains the following three directories /csjs, /css and /data. For more details regarding the result HTML directory, see 5.4.

- **vowel file:** The vowel file can be provided by the user in order to define a vowel set that differs from the default set *a, e, i, o, u*. For a detailed definition of a vowel file see 5.3.

Example of a vowel file:

a,e,y,o

Remark: This file is only needed for transcriptions or orthographies which represent vowels with letters other than *a, e, i, o, u*.



Reduplication Finder

Reduplication Finder: Pattern

Please select the file, which contains the pattern: Select a file...
D:\afessl\Arbeit\FindRedup\Veronika\patternlist.txt

Please select the file (or a dir below), in which the reduplications should be looked for: Select a file...
D:\afessl\Arbeit\FindRedup\Veronika\rawitdawitagom.txt

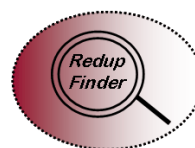
Please select a directory (or a file above), in which the reduplications should be looked for: Select a dir...

Please select the directory, where the resultfiles should be saved: Select a dir...

Please select the vowel file (default vowels are: a, e, i, o, u): Select a file...

◀ Back ▶ Next Finish Cancel

Figure 12: Step 2 after the selection of the files/directories for the pattern search.



4.2.3 Step 3 – Start Search:

Step 3 is similar for all three types of the search.

In the list search and in the pattern search, the word list file or the pattern file is displayed on top of the third window. In the free search this section is left empty, because no additional information for the search is necessary.

The currently selected search file or search directory is presented in the section “Search file/directory” below.

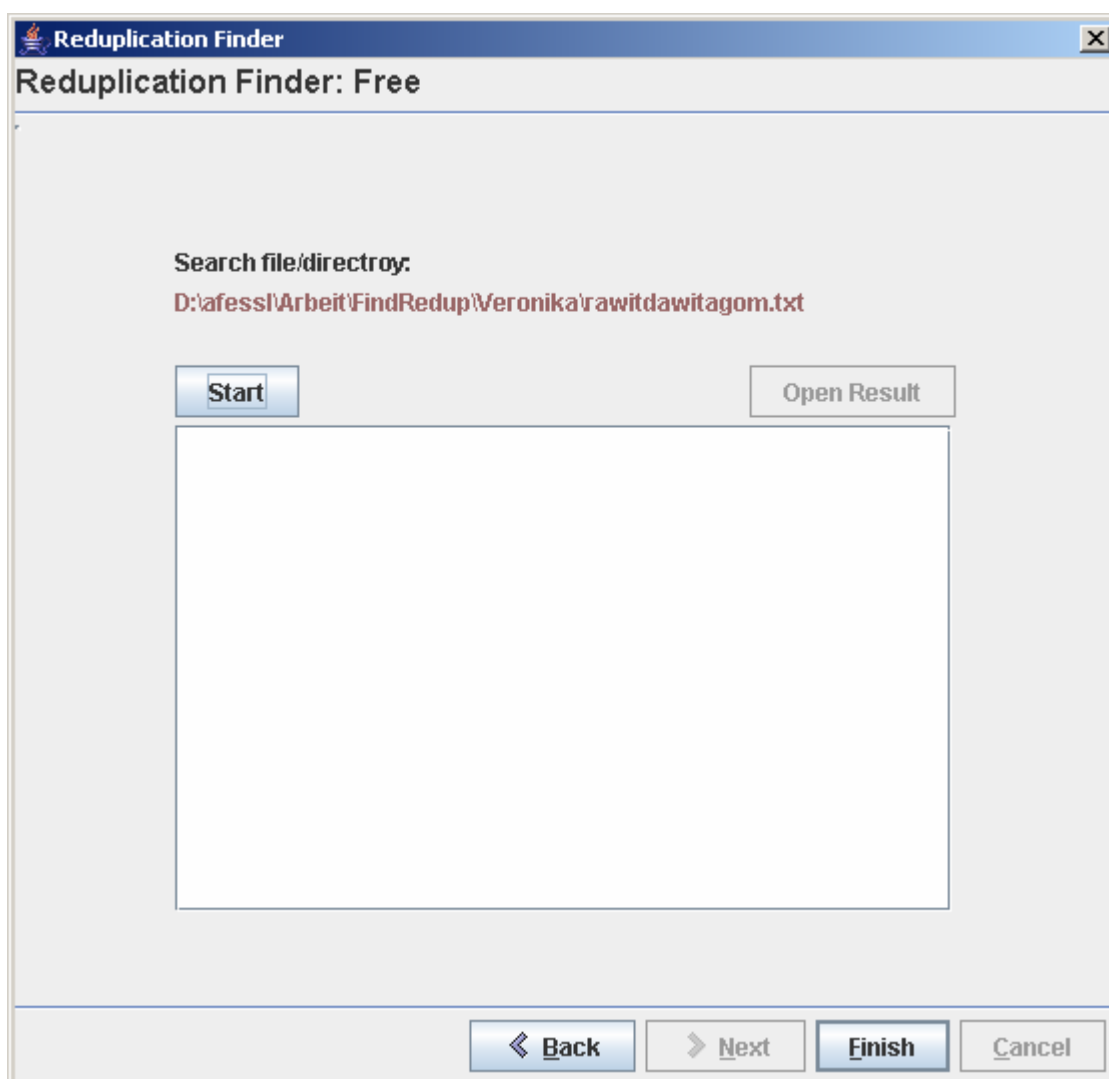


Figure 13: Step 3 for free search.

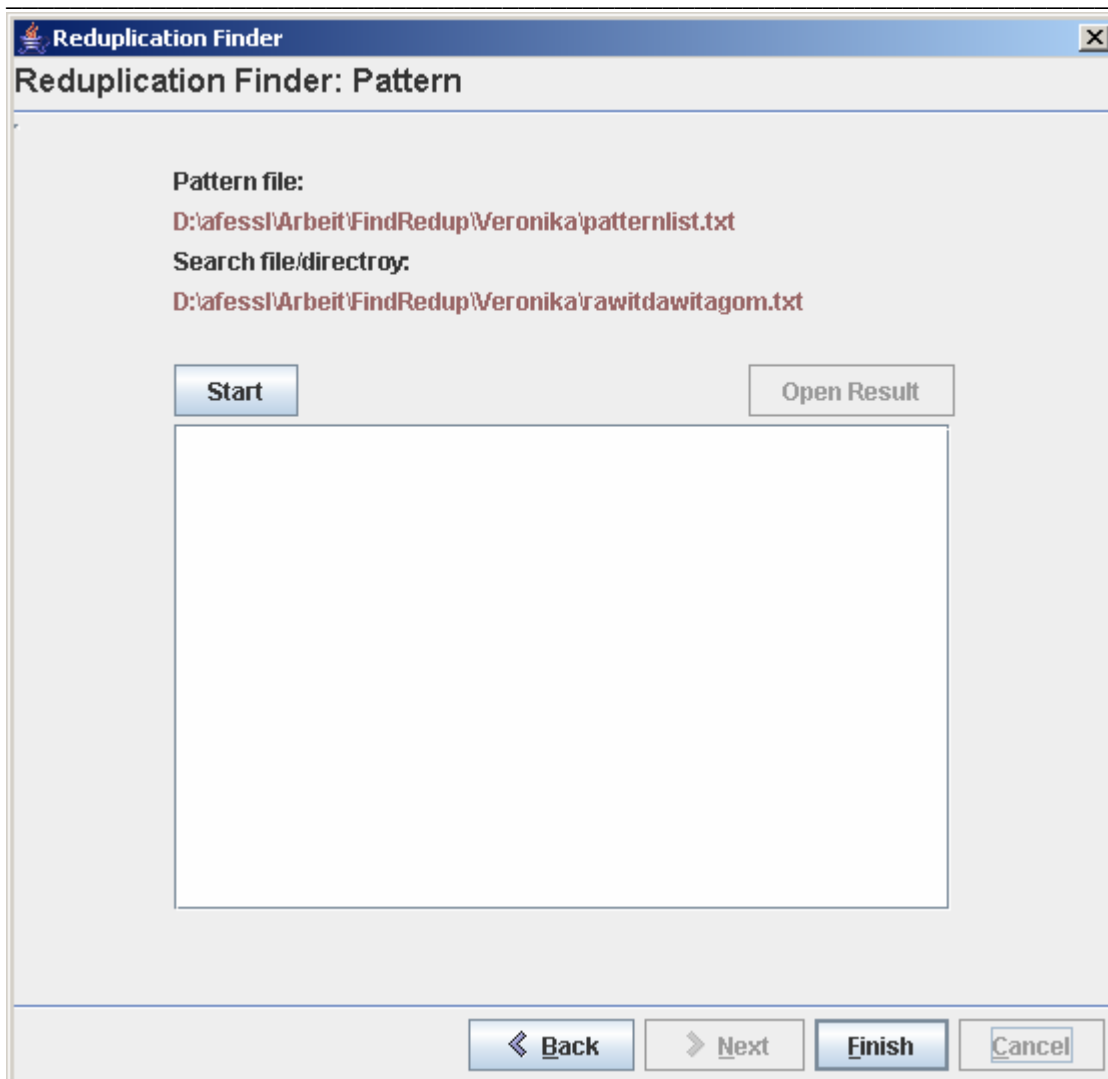
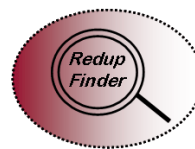


Figure 14: Step 3 for pattern search.

The “Start” button is always available:

- **Start:** To start the search, click on the “Start” button. The search starts and a progress window is opened. This window displays the current status and progress of the search. It contains the name of the file which is being searched. It also contains the number of files to be processed plus an additional one for the result HTML directory which is automatically written after all files have been searched through.

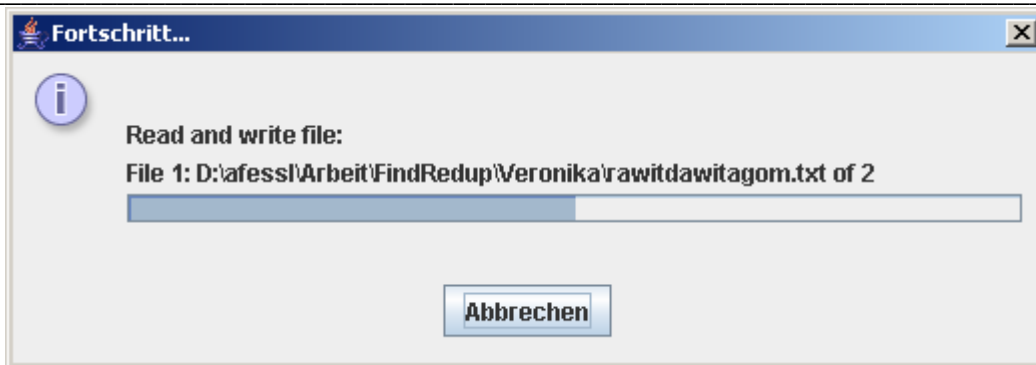
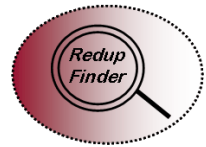


Figure 15: Progress bar after the search has been started.

- **Text area:** The program displays a text area which shows the file currently being searched. The content of the text area is updated every three seconds. When the result HTML directory is written, this is also displayed in the text area. Especially when a free search is done, it may take some time to write the result directory
- **Open Result:** the open result button will become active, when the search is finished. When clicking on this button, a browser window is opened presenting the results of the search.

After the search, the third window looks like this:

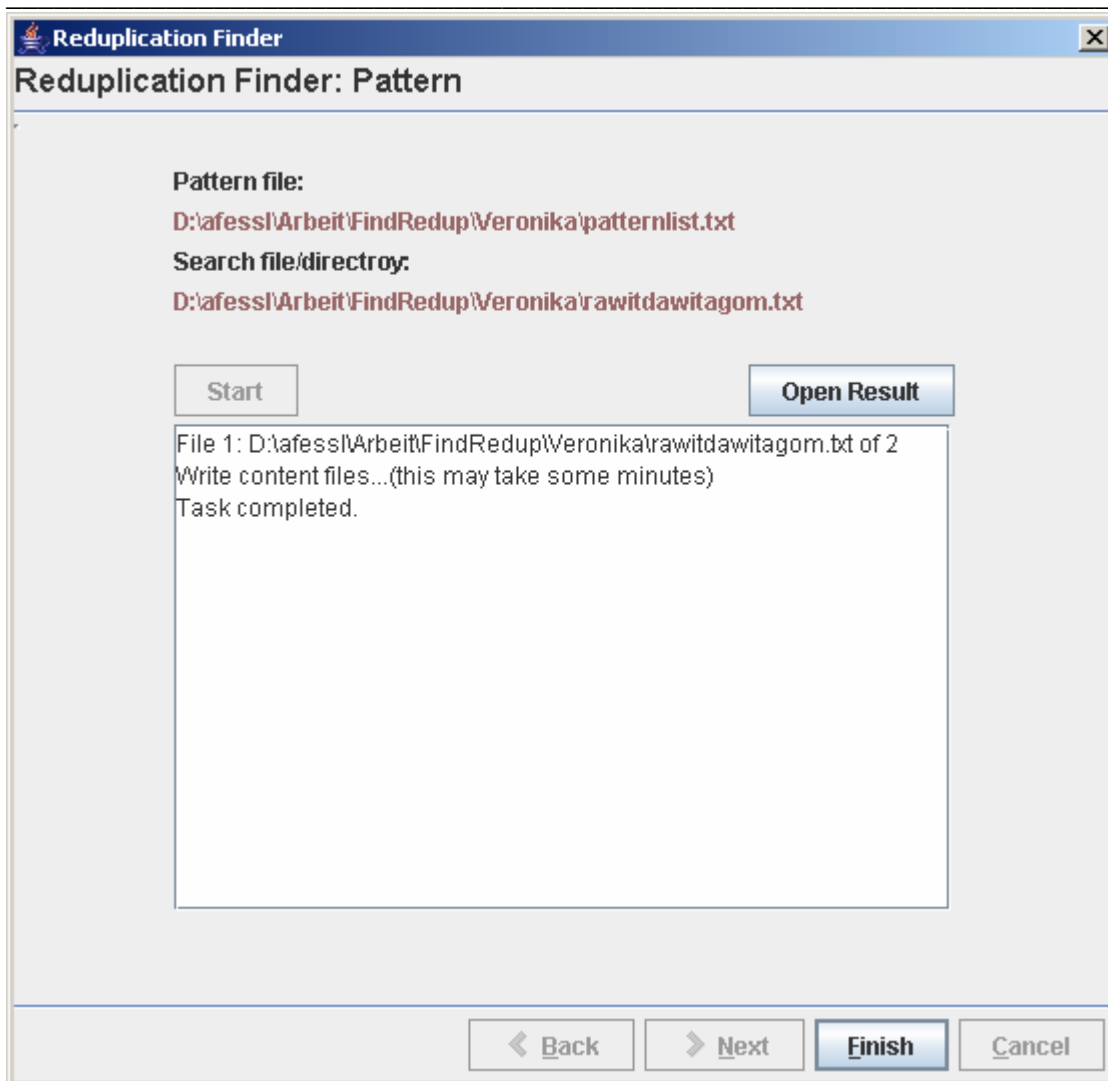
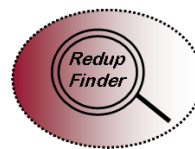
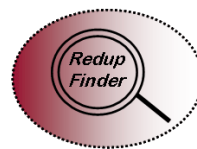


Figure 16: Step 3 after the search has finished.

After the search has been completed, click on the Open Result button to open a browser window containing the results. The user also has the option of opening the HTML directory with the explorer, which was selected in step 2. To open the result file click on the “index.html” file. For a detailed description of the html created by the Reduplication Finder, see 0.



5 Internal Description

5.1 Word list file

The word list file is provided by the user as a plain text file with the extension “.txt”. It contains a list of words (reduplications or others), to be searched for. Each word must be in a separate line. Words within a line separated by a space are treated as one word and will be searched for as well.

The word list search is case-insensitive.

For a more detailed description of how the list search works, see the example below:

Example 1:

As in all text files comment lines are allowed and start with “/” at the beginning of a line.

A word list file looks like this.:

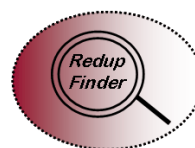
```
basa-basa
na
parapa
takig-takig
tumog-tumog
tumtom
```

An example of a content file is:

Rawit-Dawit: (Philippine poem)

```
Ki agom
(Niño Saavedra Manaog)

Nagtutururo an su'lot mong palda;
basa-basa an buhok mo, nagbuburulos
An basa sa angog mo,
Saka sa pisngi mo;
Nagtatakig-takig an ngabil mo;
Mari digdi nagparasain ka, Ne?
Nagparapauran ka na naman pauli?
Hubaa an blusa mong tumtom
Na nin lipot kan uran.
Tukaw ka digdi;
Punasan ko an payo mo; paimbungon
ko an mga kamot mo; painiton
ko an hawak mo. Nag-aalusuos na
Si sinapna ta. Gatungan ko
An kalayo nganing saimo
Igwang bagang dai masisigbo
Maski magparapauranka pa;
Maski na magparauran ka pa.
```



The results of the example 1 search are the following. Note that the forms from the word list are highlighted.

Rawit-Dawit: (Philippine poem)

Ki agom

(Niño Saavedra Manaog)

Nagtutururo an su'lot mong palda;

basa-basa an buhok mo, nagbuburulos

An basa sa angog mo,

Saka sa pisngi mo;

Nagtatakig-takig an ngabil mo;

Mari digdi nagparasain ka, Ne?

Nagparapauran ka **na** naman pauli?

Hubaa an blusa mong **tumtom**

Na nin lipot kan uran.

Tukaw ka digdi;

Punasan ko an payo mo; paimbungon

ko an mga kamot mo; painiton

ko an hawak mo. Nag-aalusuos **na**

Si sinapna ta. Gatungan ko

An kalayo nganing saimo

Igwang bagang dai masisigbo

Maski magparapauranka pa;

Maski **na** magparapauran ka pa.

5.2 Pattern file

The pattern file is provided by the user as a plain text file with the extension “.txt”. It contains a pattern list, where each pattern is in a single line.

A pattern can only consist of the following characters:

C .. Consonant

V .. Vowel

X .. Wildcards (undefined segmental strings)

Consonant or vowels always have to be specified with a number.

e.g. C1 or V2

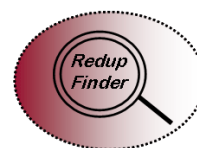
With these expressions a user can define search patterns to find reduplications. Because search patterns for different languages might be different, this type of search is very general. It is case insensitive.

If a search-pattern is defined using only C and V expressions, but without any wildcards, the search is very literal. Only whole words with the exact number of characters and with the exact order of consonants and vowels are found. See pattern examples No. 1 and 2 below.

To loosen the search for patterns add a wildcard “x”. An “x” can be put in front of or at the end of the search pattern. If an “x” is in front of a pattern, the pattern will be found at the end of the word



<http://ling.uni-graz.at/reduplication/>



searched. At the beginning of the word any characters can occur. See pattern example No. 3 below.

If "x" is at the end of the search pattern, the pattern will be found at the beginning of the word searched. After the pattern any characters can occur. See pattern example No. 4 below.

For an example with the "x" at the beginning and at the end look at pattern example No. 5.

Example for patterns and corresponding words:

No.	Pattern	Word
1.	C1V1C2V1	basa
2.	C1V1C2V1- C1V1C2V1	basa-basa
3.	xV1C1V1C2	Punasan
4.	C1V1C2V2x	kalayo
5.	x C1V1C2V1x	magparapauranka

Example 2:

This example shows a whole pattern search and how it works:

A pattern file looks like this:

```
xV1C1V1x
C1V1C2V1-C1V1C2V1
xV1V1x
xC1V1C2V2C1V1C2V2x
```

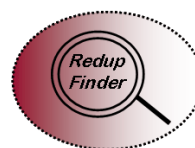
An example of a content file is:

Rawit-Dawit: (Philippine poem)

Ki agom

(Niño Saavedra Manaog)

Nagtutururo an su'lot mong palda;
 basa-basa an buhok mo, nagbuburulos
 An basa sa angog mo,
 Saka sa pisngi mo;
 Nagtatakig-takig an ngabil mo;
 Mari digdi nagparasain ka, Ne?
 Nagparapauran ka na naman pauli?
 Hubaa an blusa mong tumtom
 Na nin lipot kan uran.
 Tukaw ka digdi;
 Punasan ko an payo mo; paimbungon
 ko an mga kamot mo; painiton
 ko an hawak mo. Nag-aalusuos na
 Si sinapna ta. Gatungan ko
 An kalayo nganing saimo



Igwang bagang dai masisigbo
Maski magparapauranka pa;
Maski na magparauran ka pa.

The results of the example 2 search are the following. Note that the forms from the pattern list are highlighted.

Rawit-Dawit: (Philippine poem)

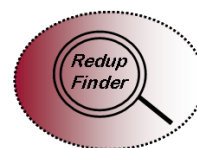
Ki agom

(Niño **Saavedra Manaog**)

Nagtutururo an su'lot mong palda;
basa-basa an buhok mo, nagbuburulos
An **basa** sa angog mo,
Saka sa pisngi mo;
Nagtatakig-takig an ngabil mo;
Mari digdi nagparasain ka, Ne?
Nagparapauran ka na **naman** pauli?
Hubaa an blusa mong tumtom
Na nin lipot kan uran.
Tukaw ka digdi;
Punasan ko an payo mo; paimbungon
ko an mga kamot mo; painiton
ko an **hawak** mo. **Nag-aalusuos** na
Si sinapna ta. Gatungan ko
An **kalayo** nganing saimo
Igwang **bagang** dai masisigbo
Maski magparapauranka pa;
Maski na magparauran ka pa.

These are the words found corresponding the given patterns in the pattern list file.

<i>Pattern</i>	<i>Word</i>
C1V1C2V1-C1V1C2V1	basa-basa
xV1C1V1x	Manaog
xV1C1V1x	Saka
xV1C1V1x	bagang
xV1C1V1x	basa
xV1C1V1x	basa-basa
xV1C1V1x	hawak
xV1C1V1x	kalayo
xV1C1V1x	naman
xV1V1x	Nag-aalusuos
xV1V1x	Saavedra



5.3 Vowel File

The vowel file provided by the user is a plain text file with the extension “.txt”. This file is only necessary for transcriptions or orthographies which represent vowels with letters other than *a, e, i, o, u*. The list of characters being defined as vowels must be in a single line, separated by a commas (“,”).

Example of a vowel file:

a,e,y,o

The program works with a pre-defined vowel set.

All characters redefined in the vowel set by the user, are not longer treated as consonants.

If a default vowel should be treated as a consonant, it is also the vowel set which must be correctly redefined. All characters which are not defined as vowels then are consequently treated as consonants.

5.4 HTML Result Directory

The results of all searches are presented in HTML. The results of the word list search and the results of the pattern search are similar, except that the results of the free search is more detailed.

The result HTML directory consists of an “index.html” file and the following three directories “/csjs”, “/css” and “/data”. Only after the free search are the additional files, “list.html”, “part.html” and “whole.html” available.

Example of the structure for the results of the list search:



Figure 17: Structure of the result after a list or pattern search

Example of the structure for the results of the free search:



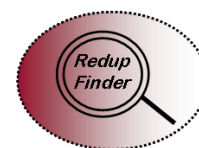


Figure 18: Structure of the result after the free search

5.4.1 Result of the free search:

The main entry point for accessing the results of the free search is an “index.html” file. The free search produces the most different kinds of results. It has the following results: list.html, part.html and whole.html.

- index.html:** The index.html file contains all reduplications found, full and/or partial. The results are ordered alphabetically. In the head of the index.html the user has the option of switching between different views of the search result: Reduplication List, Full Reduplication and Partial Reduplication. More details regarding the different results will be shown when the relevant files are described (see below).

Reduplications found:

All Words | Reduplication List | Full Reduplication | Partial Reduplication

b
[basa-basa \(1\)](#)

m
[masisigbo \(1\)](#)
[rawitdawitaqom.html](#) Preview

n
[naqbuburulos \(1\)](#)
[naqtatakiq-takiq \(1\)](#)
[naqtutururo \(1\)](#)

an hawak mo. Nag-aalusuos na
 Si sinapna ta. Gatungan ko
 An kalayo nganing saimo
 Igwang bagang dai **masisigbo**
 Maski magparapauranka pa;
 Maski na magparauran ka pa.

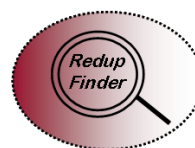
Figure 20: Index.html: Result tree view after the free search.

- list.html:** The list.html file presents only the reduplicated strings – not the entire words.

Example:

Here is a detailed result list: the patterns and the words found

A found word might be:



Nagtutururo

The reduplication found in this word is:

tutu

Thus, “*tutu*” will be displayed as a reduplication in the tree view of the list.html.

Reduplications found:

[All Words](#) | [Reduplication List](#) | [Full Reduplication](#) | [Partial Reduplication](#)

b
[basa-basa \(1\)](#)
[bubu \(1\)](#)
s
[sisi \(1\)](#)
t
[takiq-takiq \(2\)](#)
[tata \(2\)](#)
[tutu \(1\)](#)
[rawitdawitaqom.html](#) [Preview](#)
u

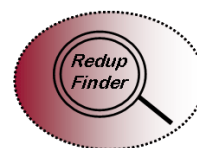
Rawit-Dawit:

Ki agom
(Nin?o Saavedra Manaog)

Nagtutururo an su'lot mong palda;
basa-basa an buhok mo, nagbuburulos
An basa sa angog mo,
Saka sa

Figure 21: List.html: Reduplication list tree view after the free search.

- **whole.html:** The whole.html shows all full reduplications found. The full reduplications may be separated by “*“*”, “*~*” or “*-*”.



Reduplications found:

[All Words](#) | [Reduplication List](#) | **[Full Reduplication](#)** | [Partial Reduplication](#)

b

[basa-basa \(1\)](#)

[rawitdawitaqom.html](#) [Preview](#)

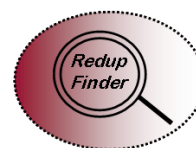
Rawit-Dawit:

Ki agom
(Nin?o Saavedra Manaog)

Nagtutururo an su'lot mong palda;
basa-basa an buhok mo, nagbuburulos
An basa sa angog mo,
Saka sa pisngi mo;
Nagtatakig-takig an ngabil

Figure 22: Whole.html: Full reduplication tree view after the free search.

- **part.html:** The partial reduplication page shows all words with partial reduplications in them.
The partial reduplications can be separated by “~” or “-“ or not.



Reduplications found:

[All Words](#) | [Reduplication List](#) | [Full Reduplication](#) | **[Partial Reduplication](#)**

m
[masisiqbo \(1\)](#)

n
[nagbuburulos \(1\)](#)
[naqtatakiq-takiq \(1\)](#)
[nagtutururo \(1\)](#)
[rawitdawitaqom.html](#) [Preview](#)

Rawit-Dawit:

Ki agom
 (Nin?o Saavedra Manaog)

Nagtutururo an su'lot mong palda;
 basa-basa an buhok mo, nagbuburulos
 An basa sa angog mo,
 Saka sa

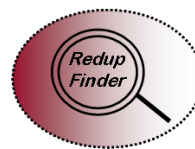
Figure 23: Part.html: Partial reduplication tree view after the free search.

- **csjs-directory:** The csjs-directory contains the same JavaScript files as the list search. Additionally it contains a JavaScript array file of the snapshots for each file searched. The name of this JavaScript array is the same as of the html-file in the data-directory, except that the extension is “.js”.
- **css-directory:** The css-directory is the same as in the list search and contains the cascading style- sheet for the design of the tree view.
- **data-directory:** The data directory is the same as in the list search and contains the searched files converted to html.

5.4.2 Results of the word list search:

The main entry point for seeing/accessing the result of the word list search is also the “index.html” file. When clicking on it the default browser is automatically opened, presenting the list of results.

- **index.html:** This file contains the result list displayable in a tree structure. The tree is alphabetically sorted and contains all entries in the word list used during the search. Access to the results is displayed initially as a list of the first characters of all words found. Clicking on a character opens the next level and shows all words in the list, starting with this character. The number of instances found is provided in parenthesis after each form. Clicking on the form displays the name of the file where the word was found. A “Preview” link also appears. Clicking on the file name opens the whole file is opened with the searched word marked. Clicking on “Preview” the immediate context in which the word



appears.

All successful results are displayed in red. All other words from the list which were not found are displayed in black.

Reduplications found:

b (1)

n (3)

na (3)

[rawitdawitaqom.html 1](#) Preview

[rawitdawitaqom.html 2](#) Preview

[rawitdawitaqom.html 3](#) Preview

p

[parapa](#)

t (1)

[takiq-takiq](#)

[tumoq-tumoq](#)

tumtom (1)

Mari digdi nagparasain ka, Ne?
 Nagparapauran ka **na** naman pauli?
 Hubaa an blusa mong **tumtom**

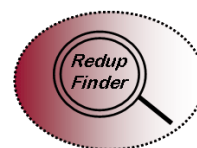
Figure 19: Result tree structure after the list search.

- **csjs-directory:** This directory contains the necessary javascript files to control the treeview, to show the snap shots and to open the window for the html content file.
- **css-directory:** The css-directory contains the style-sheet, which is responsible for the design and the colors.
- **data-directory:** In this directory the text files, which are converted to HTML during the search, are stored. Each word found in the search is marked, so that all words found can be seen at a glance.

5.4.3 Result of the pattern search

The HTML directory of the results of a pattern search is built up similar to the word list search. The main entry point is the “index.html” file.

- **index.html:** This file contains a tree view which shows at the top most level the patterns in the pattern file, where at least one corresponding word was found. The patterns are ordered alphabetically. On the next level the words which correspond with the pattern are listed. Again for each word, the file in which it was found, is shown as well as a Preview link, which presents a snapshot.



Reduplications found:

```

c1v1c2v1-c1v1c2v1
  basa-basa (1)
xv1c1v1x
  Manaog (1)
  Saka (1)
  bagang (1)
  basa (1)
  basa-basa (1)
  hawak (1)
  rawitdawitaqom.html Preview
  kalayo (1)
  naman (1)
xv1v1x
  Naq-aalusuos (1)
  rawitdawitaqom.html Preview
  Saavedra (1)

```

```

Tukaw ka digdi;
Punasan ko an payo mo; paimbungon
ko an mga kamot mo; painiton
ko an hawak mo. Nag-aalusuos na
Si sinapna ta. Gatungan ko
An kalayo nganing saimo
Igwang bagang dai masisigbo

```

Figure 24: Result tree view after pattern search

- **csjs-directory:** The csjs-directory contains the same JavaScript files as the list search and the free search. Additionally it contains a JavaScript array file of the snapshots for each file searched. The name of this JavaScript array file is the same as the html-file in the data-directory, except that the extension “.js”.
- **css-directory:** The css-directory is the same as in the list search and the free search and contains the cascading style-sheet for the design of the tree view.
- **data-directory:** The data directory is the same as in the list search and the free search and contains the searched files converted to html.

5.4.4 Modification of the HTML content

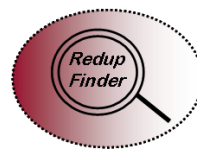
The Reduplication Finder is able to find reduplications corresponding to several rules or patterns. Of course, the program is not able to decide if a sequence of repeated segments is a **real** reduplication or only an accidental repetition of segments.

To remove a word from the index.html file, it is necessary to have a basic knowledge of html. Otherwise it is not recommended that the user attempts to edit or modify any of the files in the result directory.

How to remove an entry found in a pattern search :

The index.html file consists of 3 parts: a header, the content and a footer. For doing modifications only the content part should be changed. The other two parts should be left unchanged.

The relevant part of the html-file which the user might want to change starts with the line:



6 Copyright

Reduplication Finder 1.1

© 2006, Department of Linguistics Graz

Software Disclaimer and Copyright Statement

Permission to use, copy, modify and distribute this software for research and education purposes in academic institutions and without fee is hereby granted, provided that this copyright and permission notice appear on all copies of the software and provided that the software will not be used for commercial purposes. The “Department of Linguistics” makes no warranty or representation about the suitability of the software for any purpose. It is provided AS IS without any express or implied warranty. The “Department of Linguistics” shall not be liable for any direct, indirect, special or consequential damages resulting from the loss of use, data, or projects, whether in action of contract or tort, arising out of or in connection with the use or performance of this software. The “Department of Linguistics” cannot guarantee that the documentation associated with the software will be correct and represent the current version. The “Department of Linguistics” cannot give support for this software.

It is expected that the academic usage of this software will be acknowledged in publications. It is requested that users enter comments and errors via the Bug Report web-forms.

Questions and comments with respect to linguistic problems should be directed to reduplication@uni-graz.at.

Questions regarding support, installation of complete multimedia setups, other types of usage other than those indicated above and other aspects related to the Reduplication Finder 1.1 should be directed to: angela.fessler@uni-graz.at