

# ProefLeser 2015

## How to obtain ProefLeser

The price for one ProefLeser is R 500 (Eur 35). The price per PC/user is negotiable if more than one PC/user. Student price R250. If you would like to order the program, send us an email - [stpubus@gmail.com](mailto:stpubus@gmail.com). We will send you an invoice payable into our ABSA account or in Europe into our ING account. After payment is received we upload-download the program via **WeTransfer**. The installation of ProefLeser is very simple but all needed assistance by email is free.

Because the dictionary file is large (64 MB), it may take a few hours to download the **ProefLeser** files. Future program updates only comprise small files which will not take that long to upload-download.

Please do not give copies of ProefLeser to other people - the more we sell, the better and more frequent the improvements. Rather show it to other people and encourage them to buy it.

## Platforms

**ProefLeser** runs on Windows XP, Windows 7 and Windows 8

For us to be able to give good support, please name the folder (directory) where you will copy the **ProefLeser** files as follows:

**\_PL2015\_X86**

You can place this folder in the root (C:\), or in C:\Program Files or in C:\Program Files (x 86) or in any other folder.

## Installation

The **ProefLeser** set of files (65 MB) can be plainly copied into a folder (no typical Microsoft installation process is needed).

Once the **ProefLeser** files have been copied to your PC, you start **ProefLeser** by double clicking on **PL2015.exe**

It is a good idea to make a shortcut to **PL2015.exe** on your Windows desktop or taskbar.

## Known issues

- **ProefLeser** does not automatically detect text frames. If the user places the cursor within a text frame it will check the words within that text frame and upon reaching the end of the text frame instruct the user to place the cursor outside this text frame and continue. This is not an error in **ProefLeser** but is inherent in WinWord: try Ctrl+RightArrow (jump to the next word) within a text frame, when at the end of the text frame, it does not jump to the next word outside the text frame.
- Some special characters within words such as (optional) hyphens, non-breaking spaces (etc.) may not be handled correctly.
- Smart quotes and the Afrikaans: 'n and 't [From the Dutch *een* and *het*] may sometimes pose a problem.

## The ProefLeser interface

Note that **ProefLeser** is not part of *Microsoft Word*. It is separate program that reads and corrects a document opened in *Microsoft Word*.

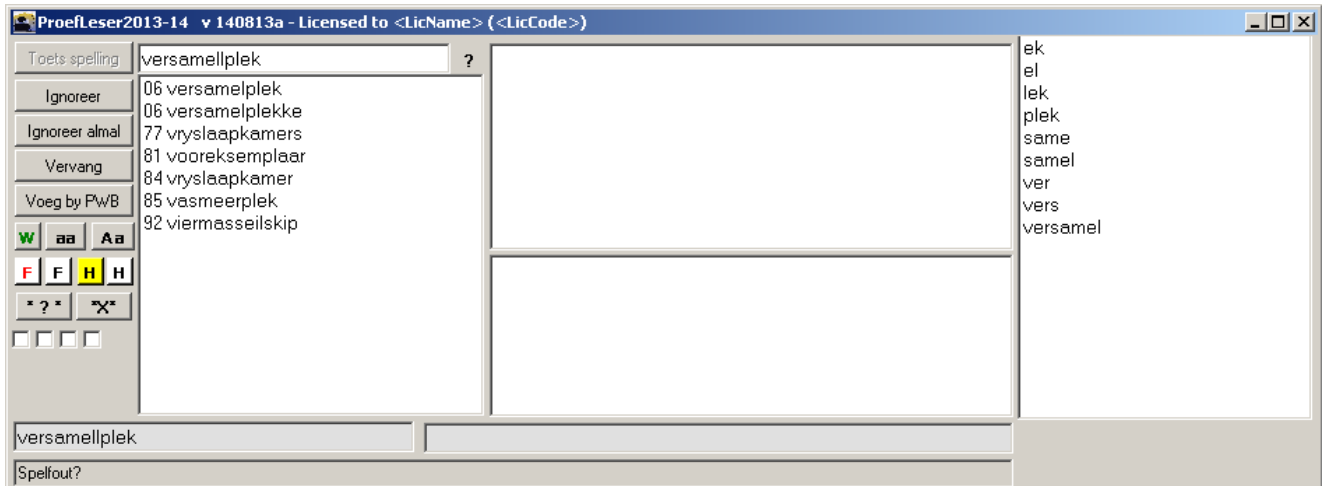


Figure 1: the buttons, textboxes, listboxes and checkboxes of **ProefLeser**

**NB ProefLeser 2015** has six checkboxes, not 4.

**NB ProefLeser 2015**, the numbers on the left of the suggested words disappear after a few seconds

- |                                     |  |
|-------------------------------------|--|
| <b>Toets spelling</b>               | Starts the spelling check process.   |
| <b>Ignoreer</b>                     | Accepts this unknown word once (case sensitive) and proceeds with the spelling process.  |
| <b>Ignoreer almal</b>               | Accepts the unknown word (case sensitive) and further occurrences of the same word. This applies only to this document.  |
| <b>Vervang</b>                      | Replaces the selected unknown word by the corrected word. You can either type this yourself or select from the suggestions. In figure 1 you would select “06 versamelleplek”. Click on “06 versamelleplek” and then click on <b>Vervang</b> .  |
| <b>Voeg by PWB</b>                  | Stores the new unknown word in the personal dictionary (PersonalDict.txt). The word will be accepted from then on, and also in other documents. You can open PersonalDict.txt with NotePad to take out unwanted words when you are not using <b>ProefLeser</b> .   |
| <b>W</b>                            | Clicking this, shifts the focus to <b>WinWord</b> . It will also enable all the above 5 buttons (some could be disabled depending on the phase/status of spellchecking). If you load or switch to another document while proofreading, click on the <b>W</b> - otherwise ProefLeser could become “confused”. |
| <b>aa</b> and <b>Aa</b>             | Changes the case in the textbox that contains the word under consideration. In figure 1 - <i>versamelleplek</i>  |
| <b>F F H H</b>                      | Will set the current selection in <b>WinWord</b> : <b>Red font</b> , Black Font, <b>Yellow highlight</b> , and/or No highlight. The idea is that during spellchecking you can mark words so that you can attend to these after using <b>ProefLeser</b> .   |
| checkbox 1 <input type="checkbox"/> | If this is switched on, <b>ProefLeser</b> will write its findings to three different text files (KumAanvaar.txt, KumOnbekend.txt and KumSamestelling.txt) for later inspection by the user. Default = off.   |
| checkbox 2 <input type="checkbox"/> | If this is switched on, <b>ProefLeser</b> will NOT stop at unknown words. This is only useful in combination with checkbox 1, i.e., it will list the unknown words in the  |

a text file KumOnbekend.txt. Default = off.

checkbox 3  If this is switched on, **ProefLeser** will stop at unknown words but will not offer suggestions.(slightly speeds up the process). You can still manually type in the corrected word. Default = off.

checkbox 4  If this is switched on, **ProefLeser** will list possible errors regarding compound and hyphenated words. This will slow down the spell-check process. Default = off.

If this is used with checkbox 1 the list of possible compound and hyphenated words will also be output to KumSamestelling.txt. This is very useful.

If switched on, you can see the running pre-context when stopping at an error.

checkbox 5  (not shown in pictures) If this is switched on, **ProefLeser** will display all the valid sub-words. Default is off – i.e. it does not show sub-words.

checkbox 6  (not shown in pictures) If this is switched off, ProefLeser keeps the focus on the program itself. If on, the focus is on MSW and the running word selection will be visible THIS DOES NOT WORK YET.

### The list and textboxes:

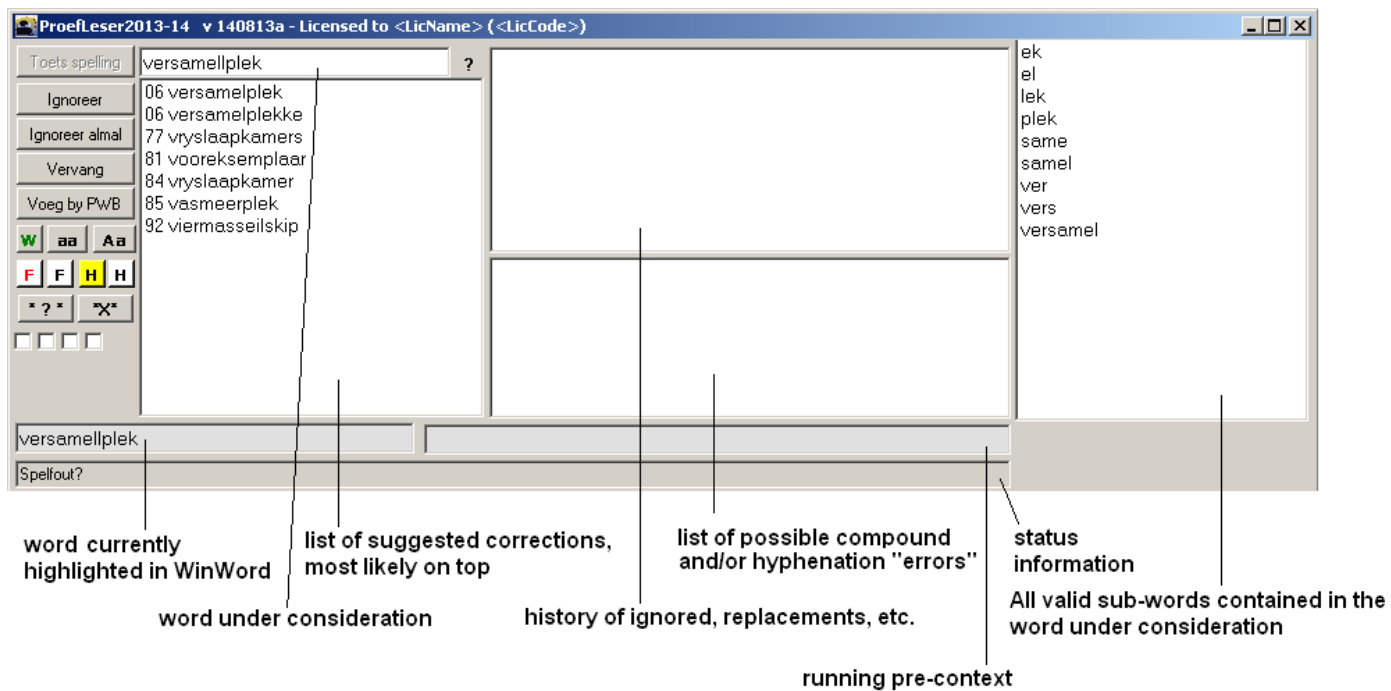


Figure 2: explanation of the list and textboxes

## Running and using ProefLeser

First load *Microsoft Word (97, 2003 ... 2010)* with your Afrikaans document. Then load **ProefLeser**. We will use the term *WinWord* to denote *Microsoft Word*.

Below is the very ancient *WinWord 97* - but all aspects also apply to later versions of *WinWord*.

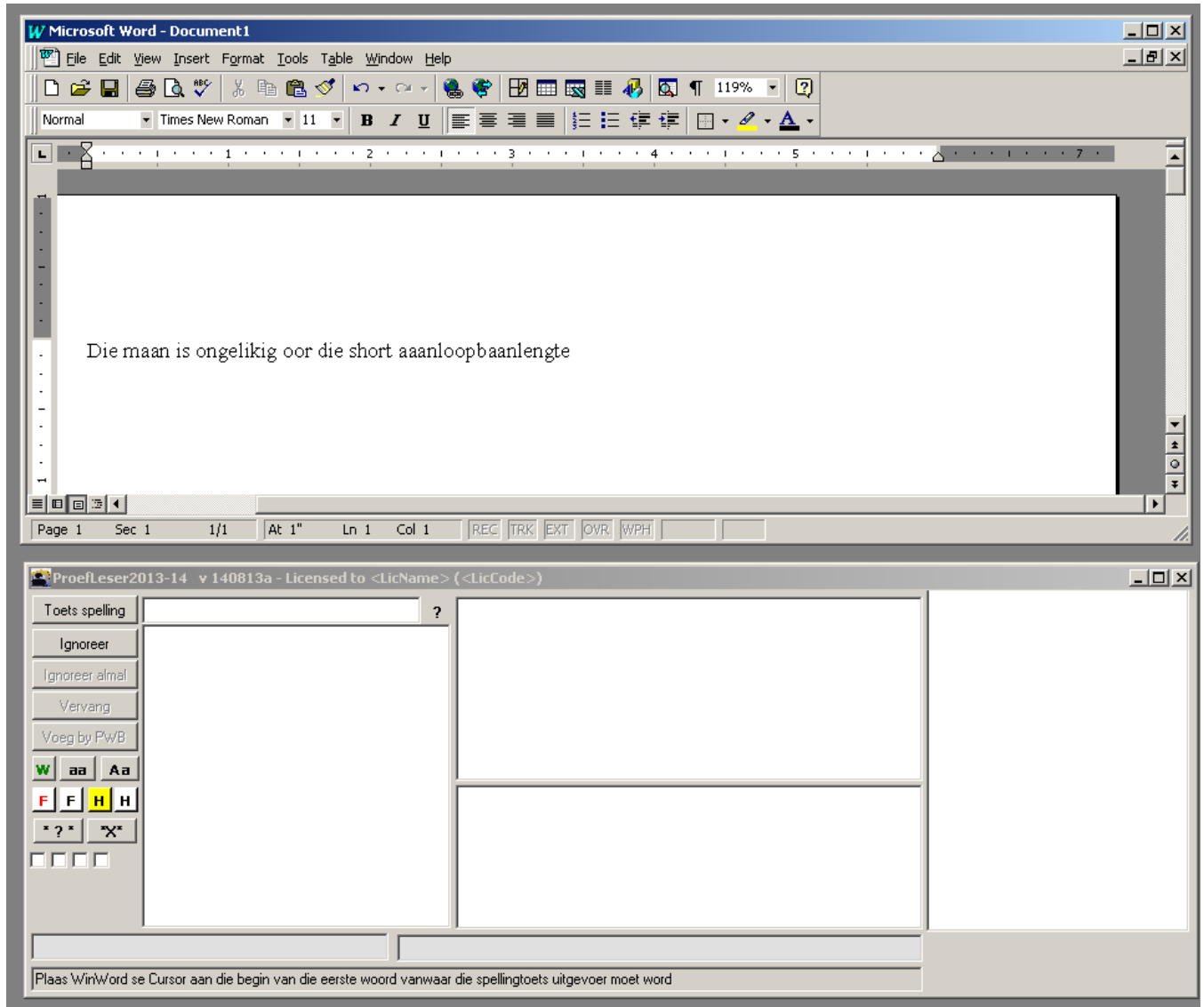


Figure 3: WinWord and ProefLeser

Click on **Toets spelling**

**ProefLeser** will start at the current cursor position in *WinWord* (it would be silly to place the cursor in the middle of a word).

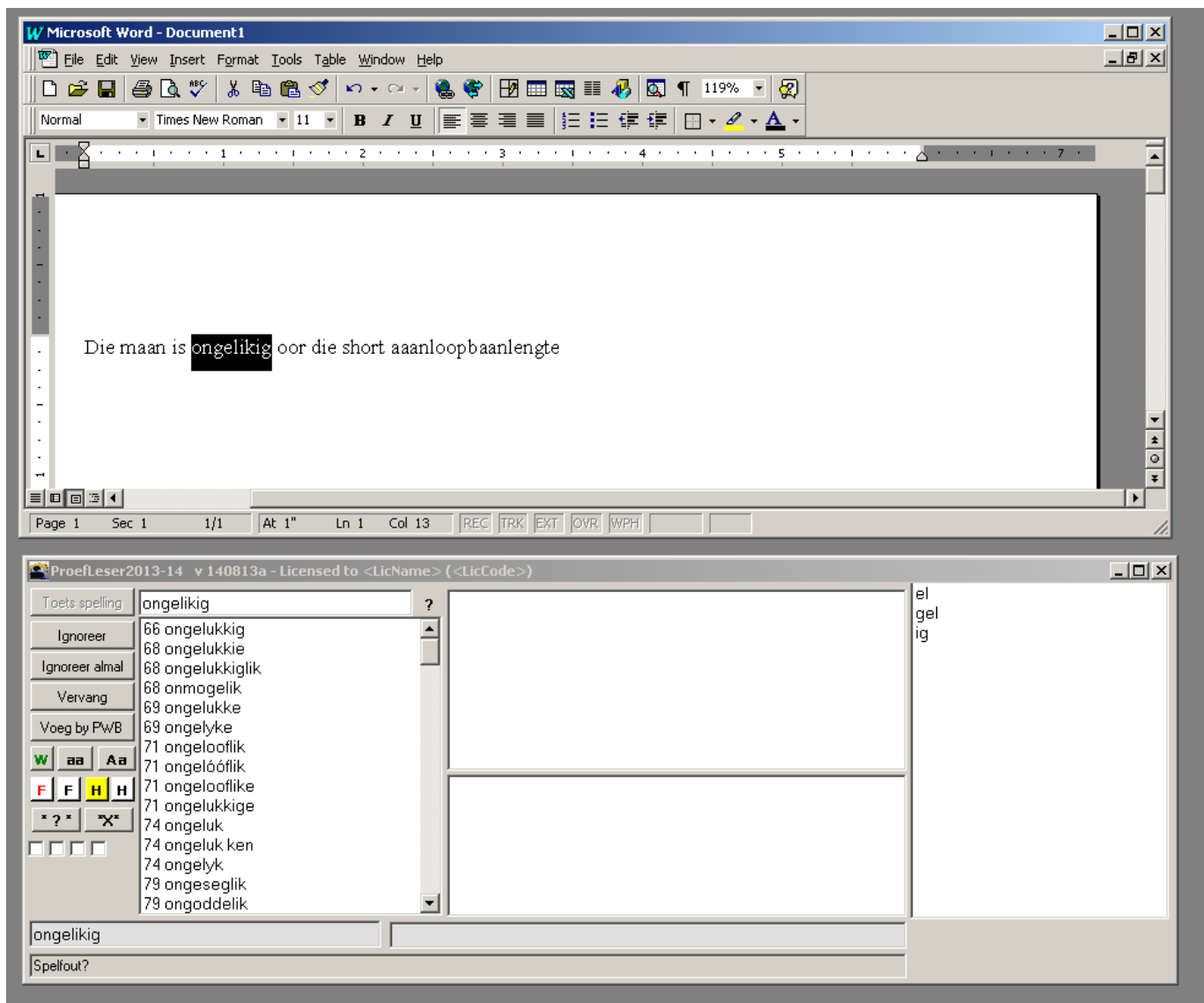


Figure 4: **ProefLeser** stopped at the first unknown word and offers (too?) many suggestions

**ProefLeser** started reading from the cursor position in *WinWord* and stopped at the first unknown word.

Obviously **ProefLeser** skipped “*maan*” although from the context it is clear that it should have been “*man*”.

It stops at “*ongelukkig*” and lists the probable correct words (suggestions). The preceding number indicates how far the correction differs from the spelling error (on a scale of 00 to 99). Clicking on 66 *ongelukkig* is fine because the 66 will not go to *WinWord*.

NB **ProefLeser** 2005 does not show these numbers but the most likely correct words are still on top.

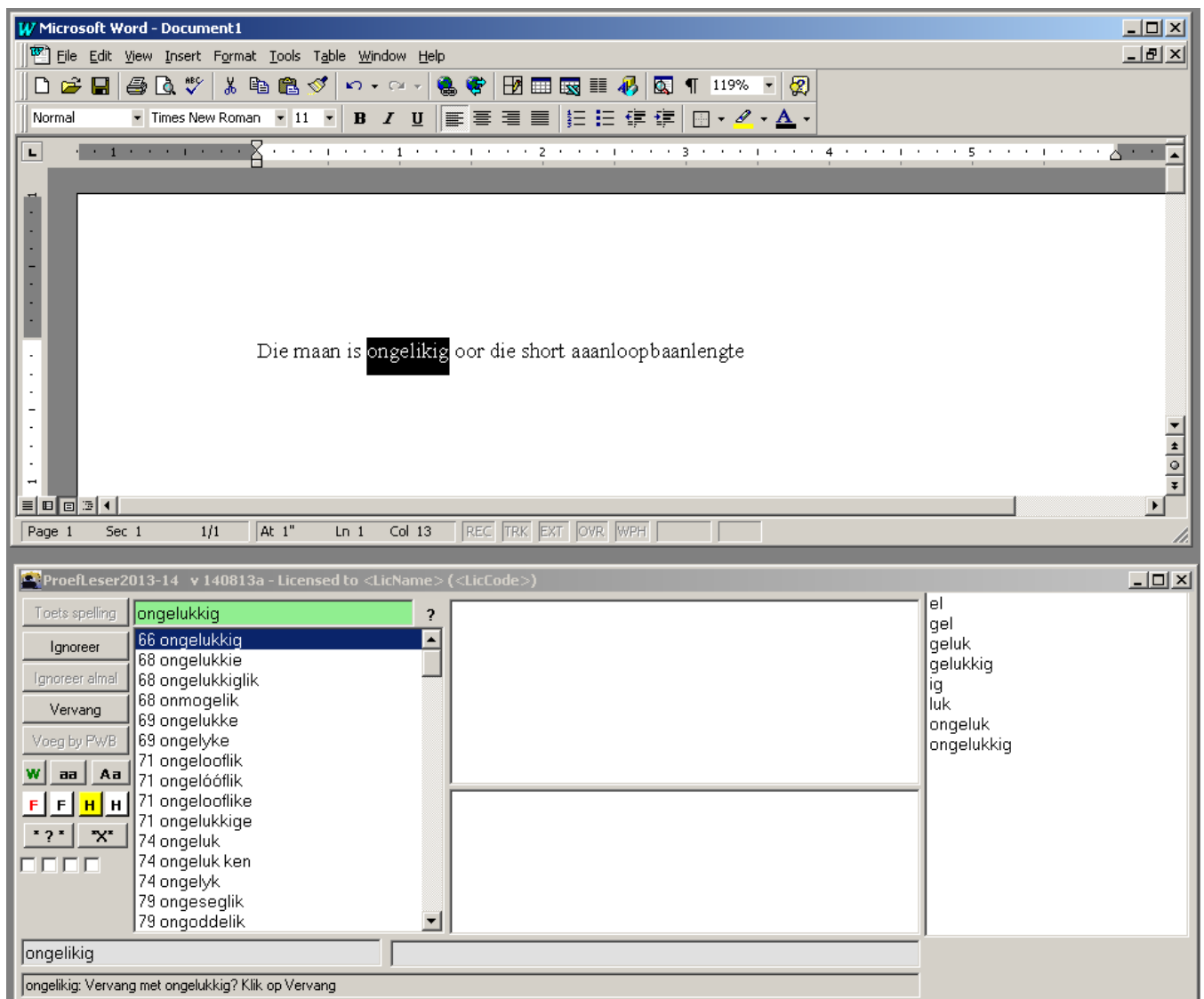


Figure 5: clicking on the proper suggestion takes it to the “word under consideration” textbox

As soon as a suggested word is chosen or an Afrikaans word that is in the main or personal dictionary is typed in, the textbox turns **green**. This means that the word is known. Note that the **grey** textbox below, still displays the highlighted word in **WinWord**.

You should now click on **Vervang**.

Note: **ProefLeser** will as far as possible logically enable and disable the 5 top buttons depending on the status/phase. For instance when it stops at an error, you have to **Ignoreer** or **Ignoreer almal** or you have to manually correct and **Vervang**. Clicking on the **W** enables all five buttons so that you are still in full control if the need arises.

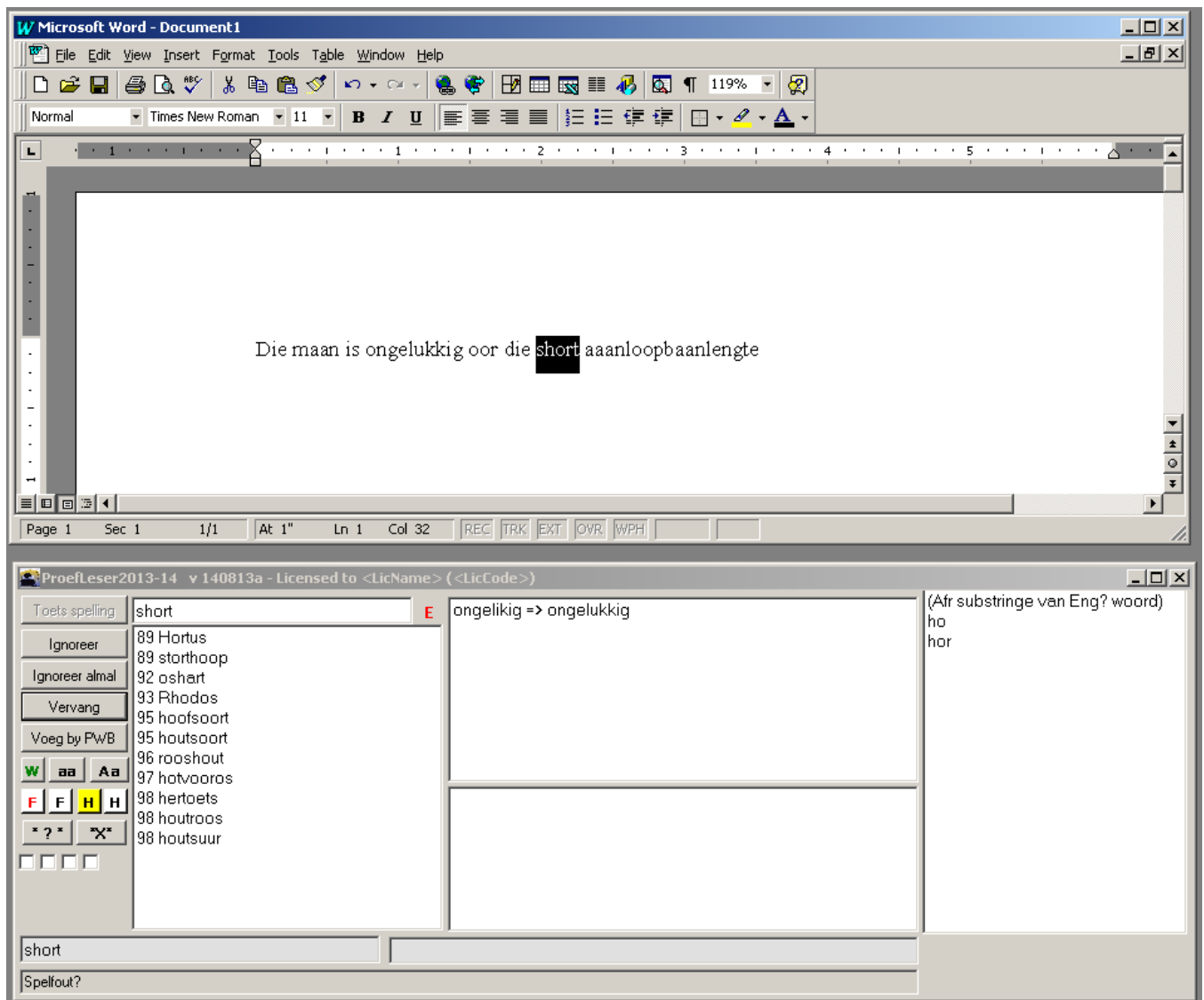


Figure 6: spelling error corrected, next unknown word encountered

The spelling error has now been corrected in **WinWord** and **ProefLeser** then stops again at the next unknown word “*short*”. It does however indicate that this word happens to be English (see the E).

Note that many words (e.g. “die”) are both English and Afrikaans.

The actions listbox (history) displays all the replacements and ignored words thus far. If you right-click on the action list you can clear it. The same applies to the list of suggested words and the list of sub-words.

## NB

Every time **ProefLeser** is loaded from scratch it forgets all the **Ignoreer (almal)** words. To remember a particular word, use **Voeg by PWB** – personal dictionary.

## NB

In the example of figure 6, if there is another occurrence of the misspelled “*ongelikig*” you can double click on the “*ongelikig* => *ongelukkig*” line in the history action list. This will put “*ongelukkig*” in the word under consideration textbox. This method allows you to consistently correct the same spelling errors.

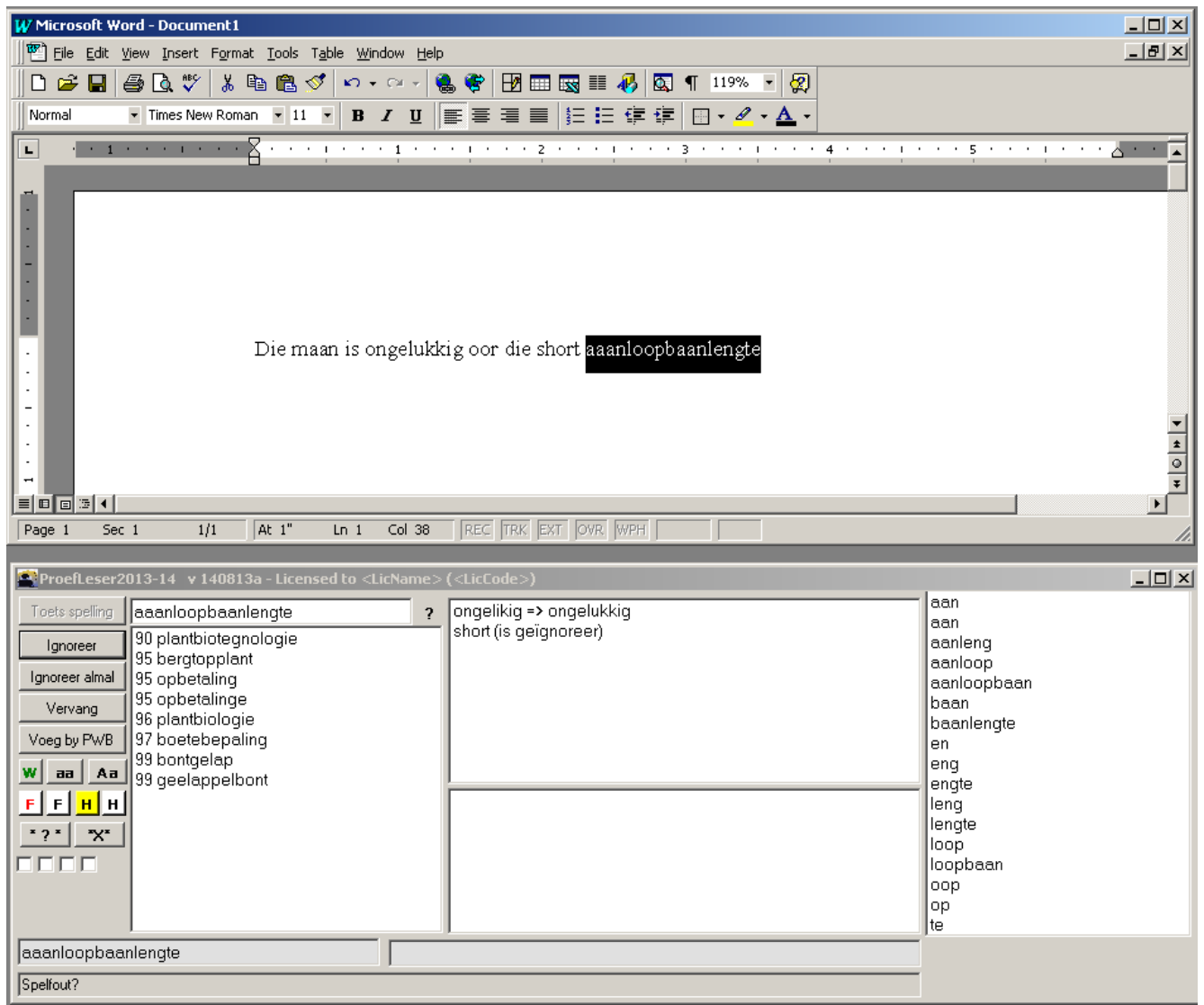


Figure 7: Compound words (“samestellings”)

After “short” is ignored (Ignoreer) **ProefLeser** will again stop at the next occurrence of “short” - if there was one. If you had clicked on Ignoreer almal it will not. It will however stop at “Short” because the case is different.

Note that the history of actions list now contains two actions.

*aaanloopbaanlengte* is also spelled incorrectly but note the listbox on the right with all the sub-words! These are **all** correct Afrikaans sub-words. It is clear that it should be *aanloopbaan* + *lengte*, so “*aanloopbaanlengte*” would be the correct word.

Note that *aan* + *loopbaanlengte* as well as *aanloop* + *baanlengte* would also make grammatical sense but semantically (meaning) less so.

You should now delete one “a” from the “aaan”. The corrected word may not be in the main dictionary (the fact that it did not turn **green** indicates that it is not) but you can add the word to the personal dictionary if you wish, and then correct it in **WinWord**.



## **Ending ProefLeser**

**Always terminate ProefLeser before unloading the document and WinWord.**

(poor **ProefLeser** keeps close contact with your document and WinWord and does not appreciate it if these suddenly evaporate).

## Advanced search functions

**ProefLeser** stops at unknown words (spelling errors) and supplies a list of suggestions. These suggestions are the words from the dictionary that are ‘almost like’ the unknown word - the ‘most likely’ ones appear on top. In addition, **ProefLeser** offers two types of wider searching methods for the ‘correct’ word. These are **WordPredictor** and searching for words that start with or end on certain characters (letters).

### WordPredictor

In the example below, ProefLeser stopped at an unknown word “fluviodglasiaal”. The user clicked on the **\*X\*** button. This opened the **WoordVoorspeller** screen containing this word “fluvioglasiaal”. The user then switched **off** the **Slegs outovoltooing vanaf die linkerkant** and removed the “d”. As the user changes the letters in this textbox, the **WoordVoorspeller** screen will dynamically display words that are ‘roughly’ similar (this is a slow process - be patient!). The user can then select a specific word from the list in **WoordVoorspeller** and place it in the word under consideration textbox with **Kies woord**.

Dynamic searching is faster if **Slegs outovoltooing vanaf die linkerkant** is on (default).

You can even restrict it to words with a specified length. Crossword puzzles spoiled!

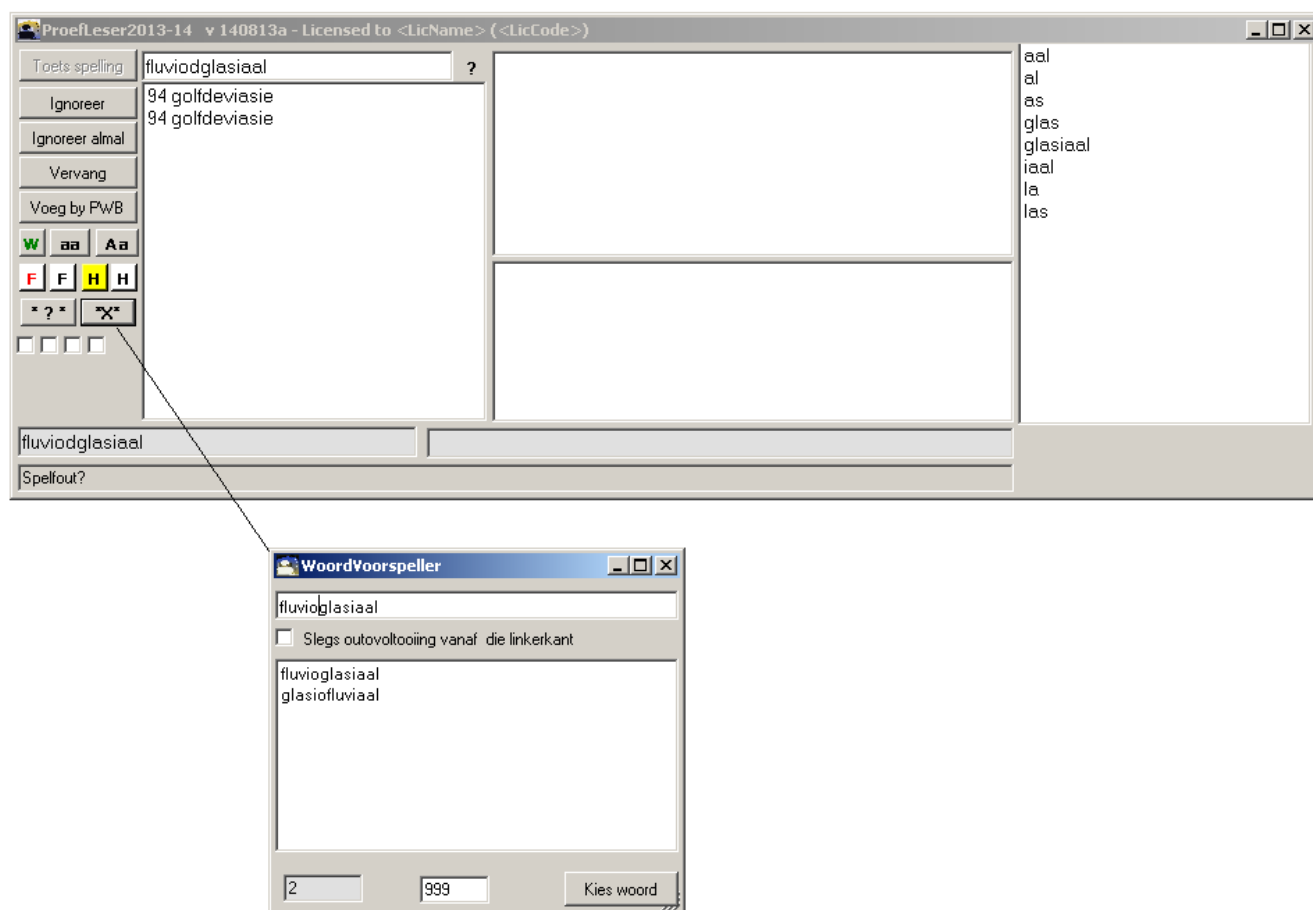


Figure 8: Searching very wide

The above example is extreme. The unknown word *fluviodglasiaal* is gibberish. The suggestions are based on words that roughly contain the same letters. The user in this case was a Geologist and he/she knew that the word did not contain a “d”. So he/she took it out and Voila! Both of the suggested words exist, innigare in fact in the *Woordeboek van die Afrikaanse Taal* - **WAT**, Stellenbosch.

## Words starting with (or ending on) certain characters

You can also directly search for all words in the dictionary starting with or ending on certain characters (limitation on number of results).

In the example below, **ProefLeser** stopped at “*knorfiets*”. Note that the option of not supplying suggestions was switched ON - for no particular reason. On the right you can see all the sub-words of “*knorfiets*”. Notice that “*knor*” and “*fiets*” are both in the dictionary, so “*knor+fiets*” could make sense. On the other hand “*knorf+iets*” would not.

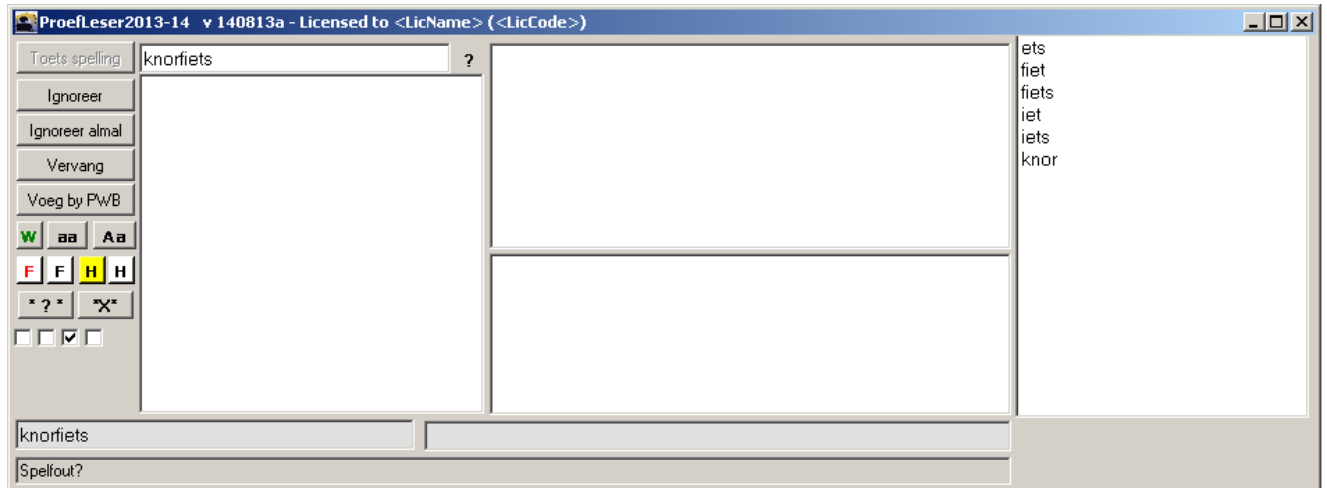


Figure 9: Looking up words

Deleting the “*fiets*” part and inserting a % after “*knor*” and then clicking on  \* ? \* supplies a list of words starting with “*knor*”. “*knortjor*” is a nice one (not visible in the figure because the user will have to scroll down).

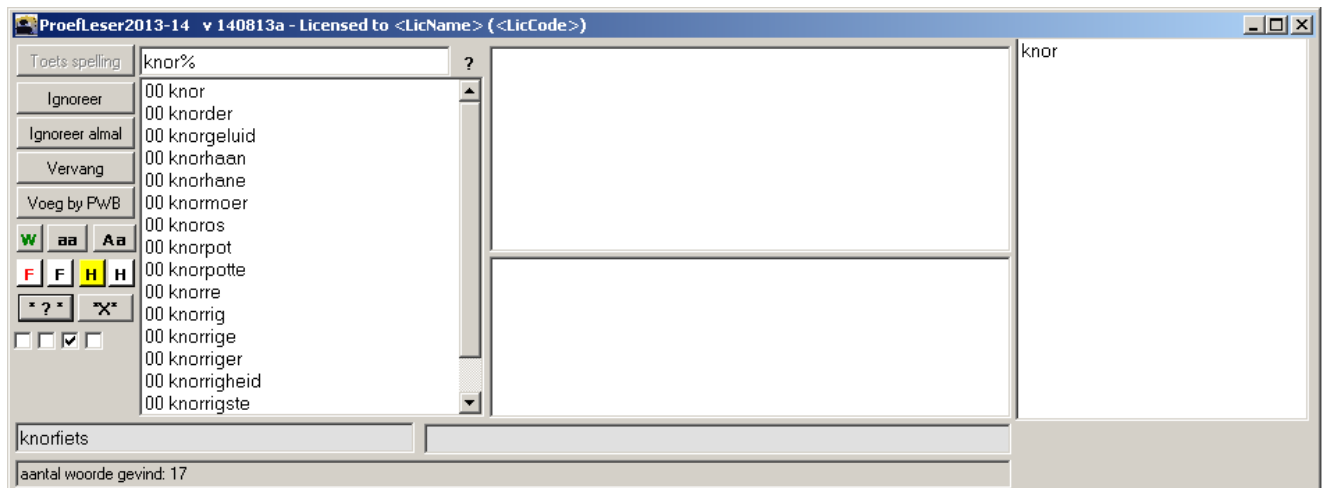


Figure 10: postfix-searching

Typing in “*%fiets*” and then clicking on  \* ? \* supplies a list of words in the dictionary that end on “*fiets*”

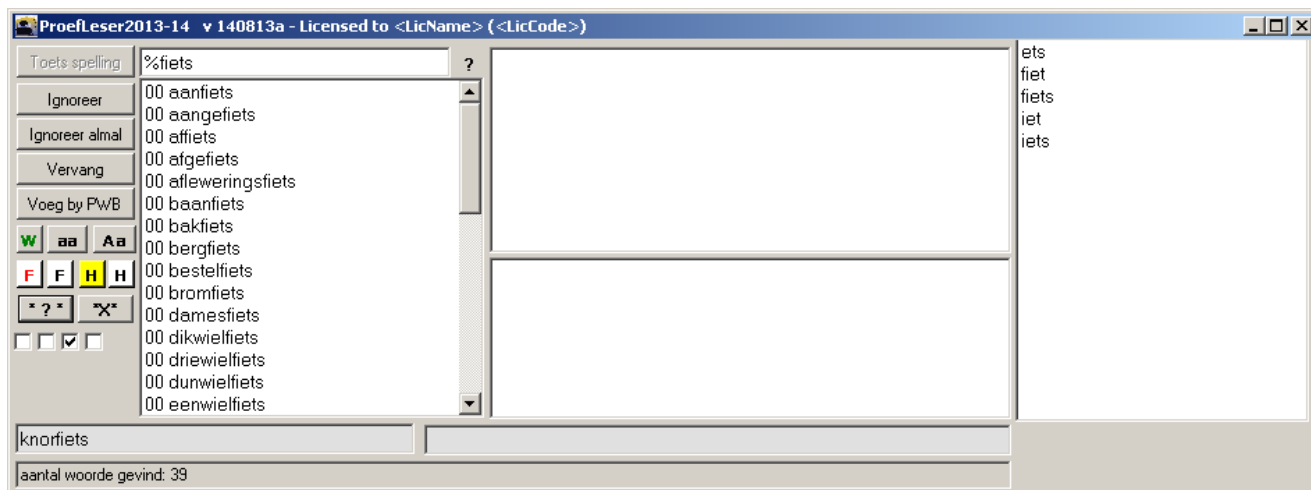


Figure 11: prefix-searching

This example shows various types of bicycles in Afrikaans.

## Compound words and other constructions

In Afrikaans there are constructions such as:

konfytekseltentoonstellingstent	- makes sense
konfytregethoutglasraam	- makes no sense but all sub-words are valid
drie en dertig	- could have a certain meaning
drie-en-dertig	- is a number
X-straal	- X is not a word
Alfa-straling	- Alfa is a word
eeue-oue	- just a nice word (only vowels)
haar fyn	- haarfyn is also a word
het sy	- hetsy ook 'n woord
hart- en vaatsiektes	- hartsiektes should also be checked
hard- en vaatsiektes	- hardsiektes should also be checked (incorrect)
20ste-eeuse	- word contains digits and "ste" is not a word
A4-papier	- A4 is not a word
as't ware of as 't ware	- Dutch: als <u>het</u> ware
om't ewe <u>of</u> om 't ewe	- Dutch: om <u>het</u> even
saldos	- incorrect
saldo's	- correct
al's	- shortened form of "alles"
ditjies en datjies	- words correct only in this combination
nasi goreng	- words correct only in this combination
dramas	- correct
drama's	- incorrect
burotjie	- incorrect
buro'tjie	- correct
fotos	- incorrect
foto's	- correct

**ProefLeser** attempts to untangle and combine some of the types listed above and compiles a list that the user can use to decide whether to search for them in the document and take further action.

This is an optional feature - it slows down the process.

The grammatical logic of this is being improved and extended on an ongoing basis.

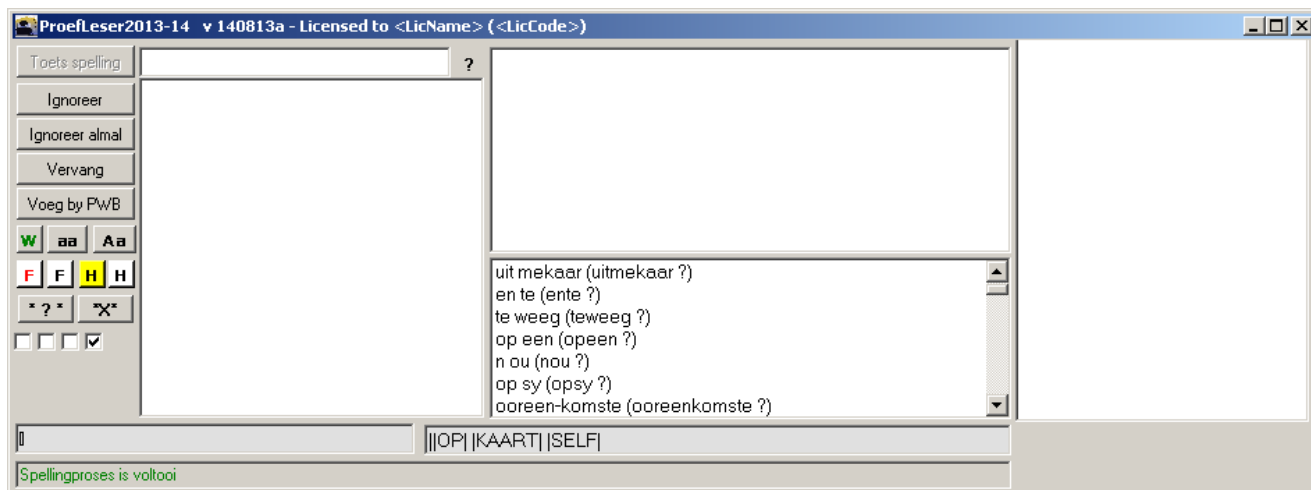


Figure 12: detecting possibly misspelled compounds (samestellings) – note that checkbox 4 is switched on

You will notice that **ProefLeser** keeps tracks of the most recent adjacent words - “[OP | |KAART| |SELF]”. In figure 12 it shows an example of a hyphenated word that could be a single word.

This list can be utilised to correct your document even further, e.g. the case of “*op sy*” and “*opsy*”:

“hy laai die pakket **op sy** wa”

“staan asseblief **opsy**”

Another interesting detection is “*en te*” versus “*ente*”

## General Hints

Do not forget to load the document with *WinWord* before starting **ProefLeser**.

Switch off the “automatic spelling as you type” (English) in *WinWord* or define the language of the whole document as Afrikaans.

If you often use the output textfiles, keep in mind that they grow and grow.

Get into the habit of arranging the **ProefLeser** and *WinWord* windows in a friendly way. Minimize all other unnecessary windows while you perform a spelling check.

**ProefLeser** always stays on top of other programs’ interfaces which means that you sometimes cannot fully see or reach other programs’ interfaces because **ProefLeser** is in the way. You can temporarily minimize and then restore *ProefLeser* again by clicking on its icon on the task bar.

## The output text files

- KumAanvaar.txt All the correct words are written to this file  
KumOnbekend.txt All the unknown words are written to this file  
KumSamestelling.txt All the dubious compound and hyphenated words are written to this file.  
PersonalDict.txt All the Voegby words are written to this file

You can open and edit these text files with *Microsoft NotePad*. Keep in mind that they should remain pure and simple text files.

You can also open and edit all the text files with *WinWord* but then make sure that you **SaveAs** them as a text file and NOT as a *WinWord* file.

**Hint:** You can load them into *WinWord* and sort them and then SaveAs a text file again.

In this way, you could assemble a useful corpus of words used by your company.

### Example of a text file

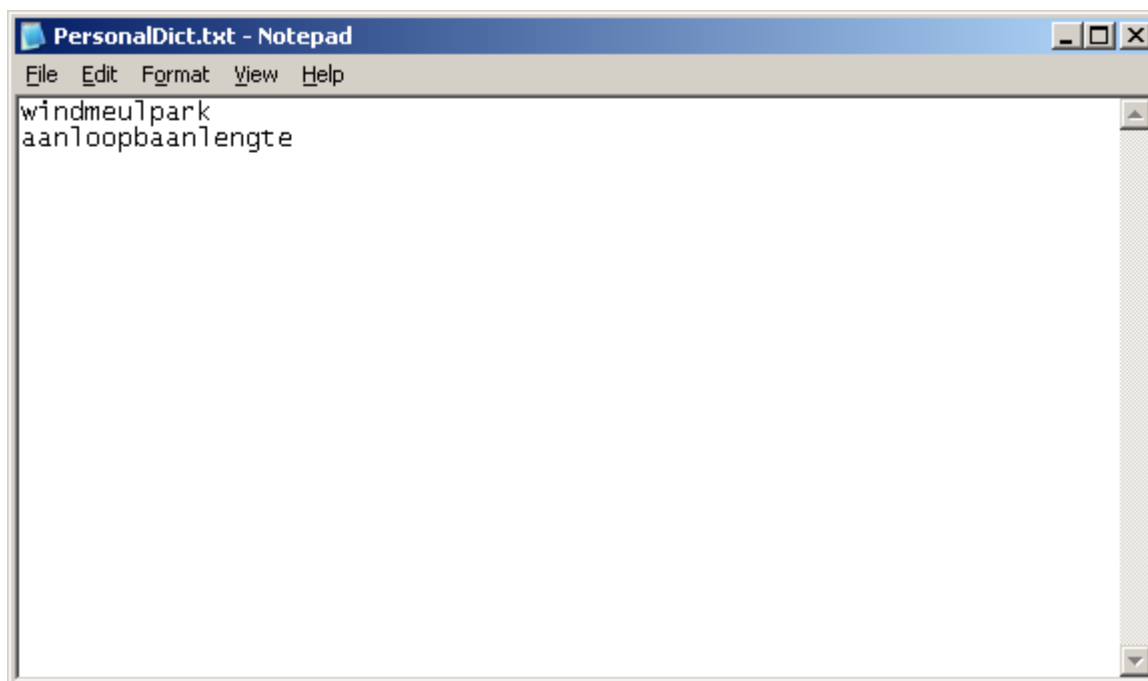


Figure 13: Text file opened in NotePad

The text file contains a straightforward list of words. Each word should be on a separate line and there should be no empty lines anywhere.

Hyphenated (e.g., see-eend) words are allowed. Short collocations (e.g., “ad hoc”) are also allowed.



## **The mathematics of finding the most likely correct word**

*Spelling checkers work but grammar checkers does not not work.*

(this part to be added at a later stage)

Last manual update:09/07/2016