# Table of Contents

What is Lingmotif? ................................................................. 3  
How does it work? ................................................................. 3  
Download and install .................................................................. 4  
  Mac OS X .................................................................................. 4  
  Windows .................................................................................... 4  
  Linux ........................................................................................ 4  
User Interface ........................................................................... 5  
  Input Modes ............................................................................... 5  
  Input Text Area ......................................................................... 6  
  Input Files Area ......................................................................... 7  
  Controls Area ............................................................................ 8  
  Other actions ............................................................................ 9  
Single-document analysis .......................................................... 11  
  Quantitative Analysis ............................................................... 11  
  Sentiment Analysis Profile ....................................................... 12  
  Detailed Analysis ...................................................................... 13  
Multi-document analysis ........................................................... 14  
  Multi-document mode – Classifier ........................................... 15  
  Multi-document mode – Series ................................................ 16  
  Multi-document – Parallel ....................................................... 17  
  Multi-file mode – Merge .......................................................... 17  
Using Plugin Lexicons ............................................................. 18  
  Creating plugin lexicons .......................................................... 18  
  Importing plugin lexicons ....................................................... 19  
  Using plugin lexicons to override Lingmotif’s core lexicon ......... 20  
  Deleting plugin lexicons ........................................................ 20  
The Sentiment Profile Tool ....................................................... 21  
  Number of points (X-Axis) ......................................................... 21  
  Y-Axis scale ............................................................................. 21  
  Saving and loading data sets ................................................... 22  
  Working with multiple profiles ............................................... 23  
About the Valence System ....................................................... 24  
  Context rules and the CVS system .......................................... 25  
  TSI Calculation ....................................................................... 26  
  TSS Calculation ...................................................................... 27  
  TSS and TSI Categories .......................................................... 27  
Publishing your results on the Web .......................................... 28  
Registering Lingmotif .............................................................. 29  
Features and Versions .............................................................. 30  
Citing Lingmotif ........................................................................ 31  
Support ..................................................................................... 31
What is Lingmotif?

Lingmotif is a Text Analytics tool, developed by the Tecnolengua Group at the University of Málaga, that provides detailed Sentiment Analysis of one or more texts. Lingmotif will assign input texts to a category on a scale, from “extremely negative” to “extremely positive”, based on a Text Sentiment Score (TSS). It also offers a measure of how sentiment-laden a text is: Text Sentiment Intensity (TSI). These data are shown graphically. A detailed, qualitative sentiment analysis is also offered.

How does it work?

Analysis is performed based on an internal sentiment lexicon of words and phrases. An input text will be analyzed and its lexical units checked against this sentiment lexicon. Lexical items in the core lexicon have different valences (negative and positive values) which will add up toward obtaining the final sentiment score (TSS) for the text.

After a lexical item (word or phrase) is identified as sentiment-carrying, its immediate context is searched for modifiers, which may change the original valence of the identified item. For example, the following sentences have different scores:

- I think he’s awesome
- I think he’s really awesome
- I think he’s extremely awesome
- I don’t think he’s awesome
- I think he’s not awesome at all

This is made possible by applying context rules to identified sentiment items. Context rules act as templates that are matched against a sentiment item’s immediate surroundings. Therefore, if a valence modifier is located outside the scope of the rule, the rule will not be matched.

Lingmotif’s core lexicon will achieve good results for most texts, but sometimes it may be necessary to override these defaults. For example, in the context of cameras or laptops reviews, the adjectives “fast” and “light” are –usually– positive, but not in other contexts. In order to account for this, Lingmotif allows you to employ user-selectable plugin lexicons, whose data will override those in the core lexicon.

---

1 If you analyze these sentences in isolation you may obtain the same TSS for those with the same polarity, but in fact they have a different score, which you can check in Advanced View (see Controls Area).
Download and install

Lingmotif is currently available for Mac OS (10.10 and above) and MS Windows (7 and above) and Linux (tested on Ubuntu 16.04). You can download the version for your platform from Lingmotif’s home page.

Mac OS X

Download the pkg file and double-click on it. Read and accept the License Agreement. Lingmotif will be installed in you Applications folder. Start Lingmotif by double-clicking this file. You can create a shortcut by dragging Lingmotif.app to the Dock.

Mac OS may complain that the app cannot be installed because it is from an unidentified developer. To override this, go to Settings – Security and Privacy – General tab – Allow apps downloaded from: Anywhere (you will need to click the padlock and enter your password). Try installing again.

Windows

Download the exe file and double-click on it. Read and accept the License Agreement. Follow the installation process.

The Windows installer allows you to install Lingmotif as “portable”. In this mode, no system files are modified, and no uninstallation is necessary. This is convenient for computers labs and other setups where regular users are not allowed to install new software. The regular installation will create a “Tecnolengua Lingmotif” entry in the Windows registry.

Linux

Lingmotif has only been tested on Ubuntu 16.04. It may or may not work on different Linux versions and GUIs.

Uncompress the tar.gz file. Place the Lingmotif folder on a (writable) location of your choice. Run the Lingmotif binary file inside that folder, either by double-clicking on it from your file manager or by typing "./Lingmotif" from a terminal.

First Run (all platforms)

When Lingmotif is started for the first time, a window will pop up asking you to provide an output folder. You need to have write permissions on this folder. If you fail to provide a suitable folder, Lingmotif will not run correctly.

On a regular setup, simply accept the suggested default (Users/username/lingmotif_output) folder. You can change this location later if you wish (see Other actions).
User Interface

Lingmotif’s main window has three sections: an Input Text area, an Input Files area, and a Controls area.

When the application starts, it will load an input test text for you to try. If you click the ‘Go’ button it will analyze that text.

Input Modes

The ‘Input’ dropdown control allows you to select the type of input for your analysis. The options are the following:
**Text box:** type in or paste text in the text box control. This is the default. It will always produce single-document analysis (see **Input Text Area**).

**File(s):** will display a dialog box that allows you to select one or more text files. Files are checked for the right encoding, type and size (according to your license type). Files that don't match the filter criteria will not be loaded. This will produce single or multi-document analysis (see **Input Files Area**).

**Folder:** this allows you to select a folder of text files. This selection is **recursive**, so it will include all files in all subfolders, too. If you don’t want to include subfolders, simply use the ‘File(s)’ input mode and select all files in the folder. This will produce a multi-document analysis if two or more files are loaded.

**Multi-Doc File:** this option allows you to **select one file only**. This file is assumed to have **one document per line**. This is a **convenient way to load a large number of short documents** (for example, Twitter text). This will produce a multi-document analysis if more than document (line) is loaded. In this mode, text is not wrapped to window width.

**Input Text Area**

You can edit the text area any way you want: select all and delete it, type your own text in, or paste any text from any other application. **This area accepts plain text only**, so whatever you paste in, only text will be kept. This means that you can copy the entire contents of a web page or text processor document, and paste it in the input text area; all non-text elements will be left out.

To clear the contents of the input text area, you can select all and delete, or click the **Clear Text Box** menu item.

The **maximum text length** you can enter depends on your Lingmotif version (See **Lingmotif Features and Versions**).

You can paste any type of data in this area; only text will be kept. Since the analysis proceeds sentence by sentence, only sentence markers are relevant.
Input Files Area

Instead of a single text, Lingmotif can analyze multiple texts in multi-document mode (see Multi-document analysis) by loading files in this area.

However, this area can also be used to conveniently load text into the Input Text area from files. In order to use this feature, do the following:

1. Select File(s) from the Input dropdown control. A dialog box will appear allowing you to navigate your file system.
2. Select the file or files you want to load. To select multiple files, hold the Control (Windows) or Command (Mac) key while clicking the desired files. Click OK.
3. Wait for the loading process to finish. This may take a while, because at this point Lingmotif analyzes your files to check they can be processed. If a file exceeds the maximum length or is the wrong format, it will be rejected.

The maximum number of files you can load at once depends on your Lingmotif version (see Lingmotif Features and Versions).

Input files should be raw text in Unicode UTF-8 encoding. Use a capable text editor to prepare your input files. We recommend TextPad for Windows or TextWrangler for Mac.

Once the files are loaded, a list will be shown in this area. Clicking any file will load it in the text area, ready to be analyzed.
Once you have clicked one file in the list you are in **single-document mode**. If you click the Go button only the selected file will be analyzed. The difference between doing this and pasting the text in the Input Text area is that the output HTML file will have as title the current file’s name (without extension) instead of the default “Text Box” title. It is also a convenient way to switch from one file to another and view their contents. You can also edit the text after it is loaded. Be aware that any changes you make will be lost after a new text is loaded (the text is loaded from the source file every time you click on it).

You can return to **multiple-document mode** by right-clicking on any element of the list and selecting “Select all files in list).

Finally, these two input areas (Input Text Box and Files List) can be **resized** by dragging the splitter that divides them. So if you are just interested in viewing/editing text in the box, you can have all the space just for this by dragging the splitter to the bottom.

### Controls Area

Here is where you select the options for your analysis and your input texts to be analyzed.

**Language**: select the language of the input text (English by default).

**Plugin lexicon**: activate a plugin lexicon to be used for the analysis. None by default. See Using Plugin Lexicons. You can have any number of plugin lexicons available, but **only one can be active at any one time**. This option is disabled in the unregistered version.

**Input**: Select the text source for the analysis. Text Box, File(s), Folder, or Multi-doc File. See Input Modes.

**Results View**: Basic (default) or Advanced. The latter will show additional data in the detailed analysis of the output. This option is disabled in the unregistered version.
Profile X-Axis: “Automatic” will let Lingmotif choose between 10 and 100 data points for the Sentiment Profile. “Fixed” forces 10 data points. This option is disabled in the unregistered version.

Profile Y-Axis: “Automatic” will make the graph show only the section of the Y-axis where values are present. “Fixed” will always show a 0-100 scale, and will add a neutrality line (at 50) for visual reference. This option is disabled in the unregistered version.

Apply CVS: CVS stands for Context Valence Shifters, and refers to the abovementioned context rules. This option should be left checked (default) unless you want to find out what sentiment valence the individual words and phrases have, regardless of their modifiers. If you uncheck this option, a text like “not good” will be analyzed as positive.

SA Items list: checking this option will include in the output a list of all text items identified in the text. They are shown in a table with three columns (positive, negative, and neutral). The list can be quite long, which is why it is unchecked by default.

SA profile: if unchecked, the Sentiment Analysis Profile will not be generated.

Go: click this button when you are ready to perform the analysis. Analysis time is proportional to text length and will also depend on your machine’s power. Once it is complete, the output will be displayed on your default web browser.

Other actions

Lingmotif’s menu gives you access to other actions and options. Be aware that some of these features are available only in the registered versions (view Lingmotif Features and Versions).

File – Open Output Dir will open a window displaying the contents of the current directory (folder) used by Lingmotif for output. This is where HTML files are generated.

File – Save XML file. For each input text, Lingmotif generates an internal XML representation that is then transformed to the HTML file that you see as output. You can choose to save the XML string as a file. Lingmotif will save the file in the current output folder and ask you whether you would like to open it in your default browser.
File – Select Output Dir. When Lingmotif is started for the first time, it will create a “lingmotif_output” folder under your home directory. You can change this default location using this option. Be aware that the old location will be left untouched. You should delete it if you don’t need its contents.

Tools – Lexicons Info. This will display information about the current lexicons in Lingmotif, both core and plugin.

View – Clear Text Box / Clear File List / Clear All. Clears the selected controls.

View – Select all files in list. Returns to series mode for Multi-document analysis.
Single-document analysis

Lingmotif’s user interface is used to customize the data sources, select the analysis input, options, and launch the analysis. The result of the analysis is generated in HTML format and sent to a new tab of your default web browser, where you can study it, search it, print it out, or publish it on a web server. In fact, you can do whatever you would do with any HTML file.

The HTML file itself is saved to Lingmotif’s output directory. When Lingmotif is run for the first time, it will create a default output directory (‘lingmotif_output’) in your home directory. You can change this location from the File – Change Output Dir menu item. Every analysis you perform will generate one or more files in this directory. In this location, you will also find other files that are required to properly show the HTML files. If you publish your results on a web site or move the files to a different location manually, you should also copy these files.

Quantitative Analysis

The two main sentiment measures of a text are provided by two quantitative measures: Text Sentiment Score (TSS) and Text Sentiment Intensity (TSI). These are displayed at the top of the page by two graphical gauges.

![Text Sentiment Score and Text Sentiment Intensity](image)

It is important to note that both the TSS and TSI measures are always on a 0 (extremely negative / factual) to 100 (extremely positive / extremely intense) scale, thus being absolute measures, regardless of text length. Text size is accounted for automatically.

**TSS** is calculated by weighing the scores of positive, negative, and neutral lexical items (grammatical words do not enter into the equation). We use the more general term *lexical items* rather than *words* because an item can be one of the following:

- An individual (non-grammatical) word: ‘excellent’ (pos), ‘death’ (neg), ‘mouse’ (neu).

---

2 If you intend to publish your analysis to a web server, it’s best to delete everything in the output folder first, then generate the analysis and simply upload everything in this folder to the server’s folder.
• A CVS segment: 'incredibly good' (pos), ‘not a good idea’ (neg), ‘absence of hostility’ (neu).

Each item is assigned a (positive, negative or neutral) score during analysis, and the final TSS is calculated based on the percentages of the accumulated scores, factoring in text size.

**TSI** is calculated as the proportion of sentiment (positive or negative) vs neutral scores. TSI itself is therefore an absolute measure itself, but it is categorized by text length. This is because the longer a text is, the more improbable it is that all its words carry sentiment.

**Text Analysis**

This section offers several stats for the input text: number of sentences, number of tokens (words), types (number of different words), type/token ratio, and number of words by function (lexical or grammatical) and by form (single words or multiwords).

**Sentiment Analysis**

This section displays all the data that Lingmotif uses to obtain the TSS and TSI measures.

**Sentiment Analysis Profile**

The SA Profile is a graphical representation of the document’s “sentiment map”. Every point in the graph represents the joint valence (normalized to 100) of the lexical items in that particular text segment. The graph below clearly represents a very positive text, since most of its points are well above zero.

Hovering the mouse pointer over a point will display a bubble with the lexical items that make up that segment, including their individual valences.
Below the main graph is a **Zoom Slider** that you can use to expand sections of the graph. This is useful when the graph contains a lot of points. Simply make the zoom slider smaller by dragging one of its ends and then drag the slider to the section you are interested in.

Lingmotif will generate automatically either **10 or 100 graph points, depending on text size**. In the registered version you can choose to force 10 segments for long texts using the X-Axis option (see **Controls Area**).

In the “series” **multi-document analysis** mode the sentiment profile represents the flow between documents. This means that that the number of points is determined by the number of documents, so the X-Axis option has no effect in the resulting graph. However, this option does have its effect on the analysis of the individual texts that make up the set.

Lingmotif’s Sentiment Profile is a powerful visualization tool. Its possibilities are described in **The Sentiment Profile Tool**.

**Detailed Analysis**

This section provides a representation of the input text marked for sentiment. Positive items are highlighted in green and negative in red. Yellow highlights indicate lexical items that were originally either positive or negative but have become neutral by the application of a context rule (CVS).

**Detailed Sentiment Analysis**

War is a state of **armed_conflict** between societies.

It is generally characterized by **extreme_collective_aggression**, **destruction**, and **usually_high_mortality**.

An **absence_of_war** is usually called "**peace**".

The registered version allows you to display the actual data for each identified lexical item by using the **Advanced View**, where each sentiment item has a superscript string in gray providing the following information:

- **VAL**: valence of the item
- **Type**: SW (single word); MW (Multiword); CVS (Context Valence Shifter)
- **Lex**: general or name of plugin lexicon (see Using Plugin Lexicons)
- **POS**: part of speech of the item. Lingmotif uses the **Penn Treebank tag set**.

War is a state of **armed_conflict** between societies.

It is generally characterized by **extreme_collective_aggression**, **destruction**, and **usually_high_mortality**.

An **absence_of_war** is usually called "**peace**".
Multi-document analysis

The Input drop-down control allows you to load and analyze multiple files at once. Multi-document mode is also activated by loading a folder (Load Dir option).

In order to enter multi-document mode, simply load more than one file and do not click on any of the files. If you do, that particular file will be loaded and single-document mode selected. This allows you to quickly select single files from a set of files or folder.

In order to return to multi-document mode, right-click on any of the file names and click Select all files in list on the contextual menu (or choose this option from the View menu). When you click the Go button in multi-document mode a dialog box will appear asking you to enter a name for the set. This name will be the title of the results page. If you don’t enter anything, the default title will be “Series/Parallel/Merged Set <current date & time>”.

When in multi-document mode, Lingmotif will analyze files one by one, generating one HTML file for each, although they will not be displayed on the browser, just saved to the output folder. There is no progress indicator while the analysis is taking place; however, you can open your file browser to the output folder and will see that HTML files are being created.

When the analysis is finished, a single results page will be displayed. This page is a summary of results and is different from the single-document results page. The gauges for TSS and TSI are now the average for the set.

![Multidocument Analysis](image)
Text Analysis

<table>
<thead>
<tr>
<th></th>
<th>Text Stats</th>
<th>Sentences</th>
<th>Lexical Words</th>
<th>Function Words</th>
<th>Single Words</th>
<th>Multiwords</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tokens</td>
<td>197196</td>
<td>6930</td>
<td>111933</td>
<td>185575</td>
<td>10916.18</td>
<td>10322</td>
</tr>
<tr>
<td>Types</td>
<td>16219</td>
<td>525.29</td>
<td>6564.29</td>
<td>85265</td>
<td>5015.59</td>
<td>607.18</td>
</tr>
<tr>
<td>T/T Ratio</td>
<td>7.93%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sentiment Analysis

<table>
<thead>
<tr>
<th></th>
<th>CVS Segments</th>
<th>Sent. Items (Totals)</th>
<th>Sent. Items (AVG)</th>
<th>AVG Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVG TSS AVG TSI</td>
<td>Total</td>
<td>AVG</td>
<td>Total</td>
<td>AVG</td>
</tr>
<tr>
<td></td>
<td>38</td>
<td>34</td>
<td>1301</td>
<td>76.53</td>
</tr>
</tbody>
</table>

In multi-document mode, the detailed analysis section contains a **quantitative analysis of each of the files** in the set. The first column in this table shows the title of the document (file name without extension) as a link to the HTML file for that particular file. You can click on it to view the single-document analysis for that file. Return to the set analysis using your browser’s ‘back’ button.

**Detailed Analysis**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>refugees: week17-2016</td>
<td>40</td>
<td>32</td>
<td>1018</td>
<td>348</td>
<td>670</td>
<td>6114</td>
<td>84</td>
<td>en</td>
<td>536</td>
<td>12728</td>
<td>2514</td>
<td>19.75%</td>
<td>12543</td>
<td>7132</td>
<td>5411</td>
<td>11717</td>
<td>742</td>
</tr>
<tr>
<td>refugees: week15-2016</td>
<td>40</td>
<td>36</td>
<td>1076</td>
<td>392</td>
<td>686</td>
<td>5450</td>
<td>85</td>
<td>en</td>
<td>533</td>
<td>12010</td>
<td>3195</td>
<td>26.60%</td>
<td>11539</td>
<td>6528</td>
<td>5011</td>
<td>10871</td>
<td>583</td>
</tr>
<tr>
<td>refugees: week16-2016</td>
<td>37</td>
<td>35</td>
<td>957</td>
<td>314</td>
<td>643</td>
<td>4921</td>
<td>65</td>
<td>en</td>
<td>486</td>
<td>10891</td>
<td>2632</td>
<td>24.17%</td>
<td>10483</td>
<td>5878</td>
<td>4605</td>
<td>9852</td>
<td>586</td>
</tr>
<tr>
<td>refugees: week14-2016</td>
<td>34</td>
<td>34</td>
<td>911</td>
<td>244</td>
<td>667</td>
<td>5049</td>
<td>82</td>
<td>en</td>
<td>471</td>
<td>10881</td>
<td>2662</td>
<td>24.46%</td>
<td>10415</td>
<td>5960</td>
<td>4455</td>
<td>9807</td>
<td>546</td>
</tr>
<tr>
<td>refugees: week13-2016</td>
<td>36</td>
<td>36</td>
<td>975</td>
<td>302</td>
<td>673</td>
<td>4872</td>
<td>78</td>
<td>en</td>
<td>474</td>
<td>10605</td>
<td>2811</td>
<td>25.51%</td>
<td>10244</td>
<td>5847</td>
<td>4397</td>
<td>9658</td>
<td>508</td>
</tr>
</tbody>
</table>

There are **four different modes to analyze multiple files**. These modes affect only the resulting **Sentiment Analysis Profile**.

The **multi-file mode** is selected in the **View – Multi-file Mode** menu.

**Multi-document mode – Classifier**

This is the default mode. In this mode no sentiment profile is generated. Instead, a stacked bar graph and data table are offered. Both offer a classification of documents based on their Text Sentiment Score. The graph offers a visualization of results; both its legend and the graph itself are interactive. Hovering the mouse pointer over the graph will show the number of documents classified in each of the categories, which are the same as the ones used in the TSS gauge.
(from “Extremely positive” to “Extremely negative”). The table shows the numerical data, and also groups results in simpler categories: negative, neutral, and positive.

### Set Classification

![Sentiment Analysis Profile](image)

Classification can be performed using ‘Multi-file’, ‘Folder’, and ‘Multi-doc File’ input modes (see Input Files Area).

Lingmotif classifies documents according to TSS categories. For comparison with binary classifiers, a coerced binary classification is also offered: in binary classification, neutral documents are classified as positive or negative randomly.

### Multi-document mode – Series

In this mode, the set of loaded files is assumed to be in order, chronological (time series) or otherwise. Each data point in the Sentiment Analysis Profile represents one document. The data point is the average TSS for that particular document. So there will be one data point for each file, regardless of the current setting in the Profile X-Axis control.

This mode allows you to discover sentiment trends and flow. For example, in order to create a sentiment timeline, do the following:

1. Place all the files for the timeline in one folder.
2. Name the files for the set so that they reflect the chronological order (e.g. “news-week01.txt”, “news-week02.txt”, etc.
3. Load the folder. Lingmotif will load the files in alphabetical order and, since it matches the chronological order, your timeline will be correct.
Multi-document – Parallel

This mode produces a graph with one dataset (line) for each file. It limits the number of files to 10, which is the maximum number of lines a graph accepts. This is useful to compare sentiment flow in texts side by side.

Multi-file mode – Merge

This mode merges all loaded individual files in one single text. The result is the same as if you saved a merged version of your files and analyzed as one file.
Using Plugin Lexicons

Sentiment is very often dependent on topic. When dealing with specialized domains, it may be useful to create your own lexicons. Lingmotif makes this task easy by allowing you to import a lexicon from a comma-separated values (CSV) file that follows a certain format.

Lexical information contained in plugin lexicons overrides Lingmotif’s core lexicon. When a plugin lexicon is selected for analysis, the plugin lexicon is searched first. If a word or phrase is found there, the core lexicon will not be searched, and the information for that item in the plugin lexicon will be used.

Creating plugin lexicons

Before importing a plugin lexicon, you need to create a CSV file. To create this file, you can use any text editor or spreadsheet software. Your file should have one or more lines with the following data structure:

word,POS,valence

Note the following:

- Commas are used to separate each of the values. No space is used after the commas.
- No spaces are to be used anywhere in the file. At the end of each line one carriage return is used.
- No first line with the name of values is used.
- Each entry is a word or multiword expression. The individual words in MWE’s are separated by an underscore (_).
- You don’t need to enter all forms for a given word. Lingmotif will automatically generate all forms for a lemma that is entered between angled brackets. For example, the following line:

  `<fast>_<speed>NN,2`

Will generate:

  fast_speed
  faster_speed
  fastest_speed

However, only one word per line will be inflected. So, if you want Lingmotif to recognize “faster speeds” as a multiword expression, you will have to enter two lines:

  `<fast>_<speed>NN,2`
  `<fast>_<speeds>NN,2`

- The **part of speech** tag is encoded using the Penn Treebank tag set.
- The **valence** for the lexical item is expressed as a positive or negative integer, or 0. Use 2, 3,... for positive items, -2, -3,... for negative items, and 0 for neutral items. Please read the section About the valence system.
It is important to understand how Lingmotif handles parts of speech. Single words and multiword expressions are treated differently:

- For **single words**, adding a POS tag to an entry means that this word should only be assigned the specified valence for that POS. For example, the line:
  
  escapes,NN,—2
  
  means that only the noun "escapes" is to be considered negative. Verbs will be neutral (unless they are marked otherwise in the core lexicon), and so will be the singular noun (if you want both to be included, you should use <escape> or add two lines).

If you want to match all parts of speech, **use the wildcard tag ALL**, as in this example:

  escapes,ALL,—2

This is convenient because you can create a list of terms and **add this tag to all of them by default** (using search and replace in your editor, or a default value in your spreadsheet). In fact, it is a good idea to use the ALL tag as much as possible for single word entries.

Finally, the **POS tag that will appear in the detailed analysis will be the one that Lingmotif’s POS tagger decided**, not the one in the lexicon. If you enter a word in the lexicon with given POS, but it is not matched during analysis, it may be due to an unmatched POS tag. Use the Advanced View to check this. If you use the ALL tag, this will never happen.

- **Multiword expressions** are not be matched against their specified tag. This is because no POS tagging is actually performed for MWE’s. So, whatever tag you use in the lexicon will be displayed in the detailed analysis (using Advanced View), but they will always be matched if the MWE is found in the text, regardless of their POS. This means that you do not really need to specify a proper POS tag, any dummy sequence of 2 or 3 characters will do: simply use the ALL tag. However, a POS tag is required; if no tag field is present, import will fail.

You do not need to include neutral items in the lexicon. Any word not in the lexicon is assumed to be neutral. Including a word with a 0 value makes sense only in two scenarios:

1. You want Lingmotif to recognize a particular sequence as a multiword expression, even if this has a neutral valence. This is particularly interesting when the MWE contains a word that would be (wrongly) tagged as having a valence.
2. You want to override a word that has a valence in the core lexicon (see Using plugin lexicons to override Lingmotif’s core lexicon).

**Importing plugin lexicons**

Once your plugin lexicon is ready to be imported, select Add Plugin Lexicon from the Tools menu. You will need to enter the following data:
• Short name: maximum ten characters (required, digits and numbers only). This is the name that will be displayed in the drop-down control and in the Advanced view of the detailed analysis section.
• Long name: a longer name for your lexicon (optional)
• Description (optional)

Lingmotif will confirm the creation of a new plugin lexicon, but no data is contained in it yet. In order to actually import your lexicon, first, select it from the “Input” drop-down control, then select Update Plugin Lexicon from the Tools menu. A confirmation dialog will be displayed.

It is important that you understand that this is the only way to modify a plugin lexicon. Once it is imported, you cannot make any changes to it (directly). However, you can modify your source (CSV) file and repeat the update process as many times as you like. So always keep your source file safe. If you lose it, you will not be able to recover it from Lingmotif.

After creating and importing a new plugin lexicon, it is ready to be used in your analysis. However, you need to activate it by selecting it in the Plugin Lexicon drop-down control.

**Using plugin lexicons to override Lingmotif’s core lexicon**

It is not possible to modify Lingmotif's core lexicon. This provides the ability to make results directly comparable: if the same tool (Lingmotif) and the same lexicon is used, even if the lexicon contains errors, results are comparable. If the lexicon is modified, this would not be possible.

However, there is a way for you to override the lexical information contained in the core lexicon. Simply create a plugin lexicon that contains the entry or entries you wish to override, and give them your desired valence (in this case it makes sense to include neutral values). Only one plugin lexicon can be used at a time, but you can have the same entries in multiple plugin lexicons.

**Deleting plugin lexicons**

You can delete any plugin lexicons you have created. Simply select it in the drop-down control and click Tools – Delete Plugin Lexicon. You can create it again at any time; just remember to keep the source CSV file.
The Sentiment Profile Tool

The sentiment Profile is a graphical representation of the text’s sentiment flow. It allows you to quickly identify where sentiment shifts occur in the text.

This section describes how you should interpret it, and teaches you how to take advantage of all its functionalities.

**Number of points (X-Axis)**

The graph is created by grouping values of identified lexical items into chunks. Each chunk is a set of lexical words (function words are not taken into consideration) and normalizing them to 100, thus always fitting in the fixed, 0-100, y-axis scale. Each point in the graph corresponds with a chunk of lexical items. Its value is the –normalized– sum of the values (valences) found in them.

However, the number of points in the graph will vary depending on a number of factors:

**Analysis mode (single vs. multi document):** in single-document mode the number of points will always be 10 or 100, depending on text size and options (see below). In multi-document mode the number of points is equal to the number of documents in the set, therefore, each point corresponds to one document. The order is the order in which the documents are loaded into the Input Files Area.

**Text size.** The number of points is determined by text size, or rather, by the number of lexical items in the text. Like so:

<table>
<thead>
<tr>
<th>Number of lexical items in text</th>
<th>&lt; 10</th>
<th>&gt;= 10</th>
<th>&gt;= 1000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of chunks / graph points</td>
<td>0</td>
<td>10</td>
<td>100</td>
</tr>
</tbody>
</table>

This means that no graph will be generated for short texts consisting of less than 10 lexical words. Also, for texts of 10 words, each point will consist of just one word, 2 points for 20, and so on. When the number of lexical items \( l \) is not a multiple of 10, the remainder \( r \) is divided amongst the last points. this means that the first \((l-r)\) points will have one lexical items less.

**User options.** For longer texts (over 1,000 lexical items) you can choose to still get 10 points using the Profile X-Axis option.

**Y-Axis scale**

The Y-Axis scale is always 100. However, by default the graph will only display the relevant section of the scale. So, if all your values are below 40, for example, the graph will only show the 0-40 section of the scale, as in the example below.
You can choose to override this behavior by selecting the **Fixed (0-100)** option in the **Profile Y-Axis** control.

### Saving and loading data sets

Since a profile is nothing but a graphical representation of a data set (the set of added valences in a particular text chunk), these values can be saved to text file. You can use these data sets to later load them and compare them. To save a profile’s data set, do follow these steps:

1. Perform an analysis where a graph is generated.
2. Click on **File – Data Series – Save Current**

The file will be saved to your “lingmotif_output” folder and you will be prompted to open the file in your default text viewer.

The name used for the file is the title of the current document. Each file will be a single line. If you plan on saving multiple profiles to later compared them you can simply copy and paste. See **Working with multiple profiles**.
Working with multiple profiles

You can work with multiple profiles in one of two ways. First, by analyzing a set of files in **parallel multi-document mode**; second, by saving the result of individual analyses and loading them prior to performing an analysis. The latter allows you to **load data sets that have not been generated by Lingmotif**.

Lingmotif allows you to load a data series representing a profile (or any other thing, for that matter) and display it along with the profile resulting of the next text analysis. You can import a saved data series from Lingmotif or use your own. An important consideration to bear in mind, however, is that the number of points in your data series must be equal to that in the **Lingmotif analysis**. So, if your data have a different number of elements, you should average them out in groups. The **format** you should use for the data series is the same used by Lingmotif when saving data series:

Name of data series=[value,value,value,…]

You can use any character for the name, but length is **limited to 50 characters**. Lingmotif, will filter out disallowed characters and shorten the name if necessary. The values should be separated by commas. The values should use the same scale as Lingmotif values, that is, 0-100.

You can load **multiple data series** from one file. Simply add as many as you need (up to 10).

Name of data series 1=[value,value,value,…]
Name of data series 2=[value,value,value,…]
Name of data series n=[value,value,value,…]
When you load the data series (File – Data series – Load from file) a message will be displayed in the status bar confirming a successful load. Click the Go button and your loaded profile(s) will be displayed along with the one resulting from the analysis. Your text will be tagged “Current Text” in the legend.

Once the graph is generated, you may use the check marks in the legend to show/hide each data series independently. You can also click on the name of a data series to hide the rest; click on it again to show them all. To stop using the data series, click File – Data series – Clear.

This feature allows you to compare data from sources other than text. For example, you might want to compare the sentiment flow in financial texts with stock market data. If the financial data are grouped by, for example, weeks (i.e., 52 data points for a year), you would group your texts per week, load your 52 text files, then load the financial data as a series, and finally perform the analysis.

Important note: When working with multiple profiles, you need to make sure that all data series have exactly the same number of data points, otherwise the graph rendering will fail and no chart will be displayed.

About the Valence System

The core lexicon and any plugin lexicons you may create share the same format and valence system. It is important that you understand how this system works.

The valence assigned to a Lingmotif lexicon entry is a number that specifies the lexical item’s polarity and intensity. This valence can be one of the following:

- A positive integer in the range 2 to 5 defines a positive lexical item.
- A negative integer in the range -2 to -5 defines a negative lexical item.
- A zero (0) defines a neutral lexical item.

You should consider these values as fuzzy quantifiers. They are supposed to capture the meaning of slightly, fairly, very, extremely.

In a nutshell, when Lingmotif analyzes a text, it identifies words and multiword expressions stored in its core lexicon and, optionally, a specified plugin lexicon, and assigns the valence specified for that word or MWE. Identified valences are added up to obtain the text’s final TSS and TSI measures.

The reason why 1 and -1 values are not used is so that items do not become neutral when a downtoning context rule is applied (see Context rules and the CVS system).
Context rules and the CVS system

CVS stands for Contextual Valence Shifters, a concept first defined by researchers Livia Polanyi and Annie Zaenen\(^3\). The basic idea is that sentiment carried by lexical items can be altered, sometimes quite significantly, by its surrounding context. For example, the valence of the adjective “great” is intensified by the presence of “absolutely” as a premodifier.

Lingmotif implements this concept in what we call context rules. Context rules are expressed as templates and searched for during analysis. If a rule is matched, the valence of the sentiment-carrying lexical item will be modified according to the rule. This is an example:

\[\text{NN},+,-,\text{prevent*},\text{LR},3,\text{INV0}\]

This rule says that a positive or negative noun, preceded or followed by a word that starts with the string “prevent” up to 3 words should have its valence inverted. It would apply to the following examples:

- The accident was prevented.
- Accident prevention is critical.
- They prevented her from winning the award.

A side effect of context rule application is that the initial sentiment-carrying item is no longer a single word or multiword item, but an item of type “CVS”. Thus, the result from applying the rule to the previous examples would be this:

- The accident was prevented.
- Accident_prevention is critical.
- They prevented her from winning the award.

As you can see, both multiword expressions and CVS segments are displayed using underscores. The Advanced View of the detailed analysis allows you to see its type.

This has an impact on TSI and TSS calculation, since the number of lexical items changes.

These are the types of shifting that Lingmotif currently contemplates:

**INVERSION.** The polarity of the original valence is inverted; e.g., INV0 -3 becomes 3. It is also possible to invert the polarity and intensify the valence:

\[\text{NN},+,-,\text{put_an_end_to},\text{L},3,\text{INV1}\]

**INTENSIFICATION.** The original value is intensified in the specified value; e.g.: INV

---

VB,+-,a_lot,R,3,INT2

**DOWNTONING.** The valence is lessened in the specified value; e.g.: **DOW1** -3 becomes -2:

```
RB,+-,less,L,1,DOW1
```

**FINAL VALUE.** The result of the rule application is a specified value, regardless of the original:

```
JJ,+-,as_thoug,L,4,VAL0
```

Context rules possess great expressive power, but they do have certain limitations, and will sometimes be applied incorrectly. For example, a modifier can be outside the scope of the rule, thus not identified, or it may actually be meant to modify another word. Even so, accuracy is generally acceptable. Another limitation in its current implementation is that rules are only applicable to sentiment-carrying words.

Context rules cannot be modified by end-users, but we appreciate suggestions and contributions. Please submit your suggestions to the Lingmotif Users Group.

**TSI Calculation**

The Text Sentiment Intensity is a measure of a text’s density of non-neutral lexical items. Therefore, the higher the proportion of sentiment-carrying words and multiword expressions, the higher the TSI.

However, this measure is not just a simple proportion calculation based on the number of items. The following should be borne in mind:

- **Items are not just counted:** since every Lingmotif lexical entry has a valence (i.e., polarity plus intensity), it is valence that is taken into account. For example, the following two sequences of valences (meant to reflect a set of lexical items) would obtain a different TSI, even though they have the same number of positive, negative, and neutral items:

  0, 0, 3, 0, 5, 4, 4, 0, -1, 4
  0, 2, 1, 0, 2, 3, 2, 0, -4, 2

- **Text size is taken into account.** Lingmotif is designed to work with any text type, from tweets to novels. When calculating the sentiment-to-neutral ratio, it would be unrealistic to simply weigh sentiment against neutral items, regardless of text size: it is highly improbable (not to say impossible) to have a 90 percent ratio of sentiment items in an essay-long text, for example. But this is quite common in one-sentence utterances (“I’m loving it”, “what a terrible day”). If text size is factored in, the resulting TSI can be offered as an absolute value on a scale (relative to text size, in fact).

- **Function words do not enter into the equation.** Function words, or grammatical words (conjunctions, prepositions, pronouns, etc.), do not add any lexical meaning to an
utterance; its presence is merely required by the grammar of the language. Such words are ignored for TSI calculation purposes. Therefore, the two utterances below would produce exactly the same TSI, even though they differ considerably in the number of “words”.

That[] was[] a[] facepalm[-3] moment[0] for[] everyone[]
Facepalm[-3] time[0]

### TSS Calculation

The Text Sentiment Score is the most important value offered by a Lingmotif analysis. It is to be understood as the global valence for a text. Lingmotif’s equation for TSS calculation uses the same parameters as TSI calculation. In fact, the TSI value is used to calculate TSS.

For calculating TSS, positive and negative values are added after factoring in text size for each value. That is, each value is weighted proportionally to text size: the longer the text, the heavier the weight. Since TSI already reflects text size, it is used to establish the weights.

This is why you can obtain an “extremely positive / negative” result from a long text with a high proportion of neutral words.

### TSS and TSI Categories

TSS and TSI are numeric values on a 0-100 scale. Although these are absolute values, in fact, they are relative to text size, as explained above.

Lingmotif also establishes ranges of values to make the numeric results more intuitively understood. Such ranges, or categories are displayed on the graphical gauges.

Below are the value ranges and categories currently used.
### TSS ranges and categories
- 0-20: extremely negative
- 21-30: very negative
- 31-40: fairly negative
- 41-47: slightly negative
- 48-53: neutral
- 54-60: slightly positive
- 61-70: fairly positive
- 71-80: very positive
- 81-100: extremely positive

### TSI ranges and categories
- 0-54: factual
- 55-59: slightly intense
- 60-74: fairly intense
- 75-84: very intense
- 85-100: extremely intense

These ranges are meant to reflect the inherently vague meaning of the quantifiers (slightly, fairly, very, etc.), which is why they differ in size.

## Publishing your results on the Web

Lingmotif generates its output in HTML/Javascript format. This means that you can have your results published on a web server straightaway. Simply upload the contents of your "Lingmotif_output" folder to a folder on your server. If you publishing a multi-file analysis, you may want to rename the main html file (the one with links to all the other documents) as "index.html", so this page is displayed automatically when the client's browser is pointed to the online folder. You must include the subfolder "css", otherwise the Javascript elements (gauges and profile) may not be displayed properly.
Registering Lingmotif

Lingmotif is **FREE FOR NON-COMMERCIAL PURPOSES**. When Lingmotif is first installed some options are disabled. In order to enable all features, you need to obtain a **STANDARD** user’s License Key. You will find instructions to obtain your key in the menu item `Help – Register Lingmotif`.

Once you have obtained your key, go to `Help – Enter License Key` and enter the required data. All options will be enabled. You only need to enter your key once, but you will need to do this again if you reinstall Lingmotif, so keep your license key safe. If you are interested in obtaining a **PROFESSIONAL license**, including use for commercial applications, [contact us](http://tecnolenguama.uma.es/lingmotif).

Your license key is personal and should only be used by you, but you may use it in more than one machine. We wish to keep track of our user base.
# Features and Versions

<table>
<thead>
<tr>
<th>Feature</th>
<th>Unregistered</th>
<th>Standard</th>
<th>Professional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max Text Size</td>
<td>3 Kb (600 words)</td>
<td>500KB (100k words)</td>
<td>8 Mb (1.6M words)</td>
</tr>
<tr>
<td>Max Files in Multi-document mode</td>
<td>100</td>
<td>300</td>
<td>10,000</td>
</tr>
<tr>
<td>SA Items List</td>
<td>✗</td>
<td>✗</td>
<td>✔</td>
</tr>
<tr>
<td>X-Axis customization</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Y-Axis customization</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>CVS option</td>
<td>✗</td>
<td>✗</td>
<td>✔</td>
</tr>
<tr>
<td>Plugin Lexicons</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Results View</td>
<td>Basic</td>
<td>Basic</td>
<td>Basic/Advanced</td>
</tr>
<tr>
<td>Save XML Files</td>
<td>✗</td>
<td>✗</td>
<td>✔</td>
</tr>
<tr>
<td>Save Data Series</td>
<td>✗</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Load Data series</td>
<td>✗</td>
<td>3 series max</td>
<td>15 series max</td>
</tr>
</tbody>
</table>
Citing Lingmotif

If you use this software for your own research, please cite it:


Support

Although no direct support is guaranteed, we will try to provide support to as many requests as possible. A Google Group has been set up where Lingmotif users can discuss issues and request improvements. We welcome each and every suggestion and bug report, but we cannot guarantee that all will be addressed immediately.

Lingmotif Users Group

Please join Lingmotif Discussion here:

http://tecnolengua.uma.es/?page_id=166

Lingmotif License Agreement

Before using Lingmotif, make sure you understand its terms of use.

Please read Lingmotif’s EULA.