

Lexus Manual

**Some parts of the manual are incomplete
because the tool is still in development.**



Lexus Manual

Jacqueline Ringersma

Some parts of the manual are incomplete because the tool is still in development.

Published LEXUS Version 2.0 RC 3 (revision r18384)

Table of Contents

1. Introduction	5
1.1. Lexus	5
1.2. Registration and login	6
1.3. LEXUS workspace	6
1.4. Workflow	7
2. Create lexica	9
2.1. New lexica	9
2.2. Defining a lexicon structure	9
2.2.1. Schema and View Editor	9
2.2.2. New elements	10
2.2.3. Modify an existing structure element and use the data category repository	10
2.2.4. Defining Word List View	11
2.2.5. Defining the Lexical Entry View	13
2.3. Lexical entries	14
2.3.1. Lexicon editor	14
2.3.2. Adding new lexical entries	15
2.3.3. Modifying lexical entries	16
2.4. Adding multi media to lexical entries	16
3. Importing and exporting lexica from other formats	19
3.1. Importing lexica from Toolbox	19
3.2. Importing lexica from XML	21
3.3. Exporting lexica to XML	22
3.4. Import Toolbox data to an existing lexicon	22
4. Searching and filters	25
4.1. Search	25
4.1.1. Ad hoc search	25
4.1.2. Creating queries	25
4.1.3. Using queries	27
4.2. Filters	27
4.2.1. Creating filters	27
4.2.2. Using filters	27
5. Characters and sorting	28
5.1. Using non-latin characters	28
5.2. Sort orders	28
5.3. Creating sort orders	28
5.4. Applying sort orders	29
6. Access rights	30
7. ViCoS	31

Chapter 1. Introduction

1.1. Lexus

LEXUS is web-based lexicon tool developed by the Technical Group at the Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands (MPI). It is targeted at linguists doing field research or working with language corpora.

In LEXUS users may create lexica in online (but private) workspaces and share lexica with other users either on a read-only basis or by assigning edit rights. Since LEXUS is web-based this sharing functionality allows a group of users to create a lexicon from different places in the world, or to share lexica amongst linguistic researchers and speech community members.

With LEXUS you may create lexica from scratch or import lexica from other formats, like Toolbox or XML. LEXUS is not just another lexicon tool: LEXUS offers the possibility to visualize language through the import of multimedia. Three different types of multimedia fragments can be linked to the lexical entries: images, videos and audio. LEXUS also interacts with media stored on the Internet. For the DoBeS projects, MPI research and for research from third parties, data is stored in the digital archive for linguistic resources housed at the MPI. The archive is accessible via the Internet, and is organized in a structured manner by describing and contextualizing the data with the IMDI metadata set. Lexical entries in LEXUS can be linked to domains and resources in this archive.

Lexicography in general is a domain where uniformity and interoperability have never been the operative words: depending on the purpose and tools used different formats, structures and terminologies are being adopted, which makes cross lexica search, merging, linking and comparison an extremely difficult task. LEXUS is also an attempt at putting an end to this problem. Being based on the Lexical Markup Framework (LMF), an abstract model for the creation of customized lexicons defined following the recommendations of the ISO/TC 37/SC 4 group on the standardization of linguistic terminology, LEXUS allows on the one hand to create purpose-specific and tailor-made lexica, and on at same time assures their comparability and interoperability with other resources.

In LEXUS, each lexicon consists of a structure and of content. Each lexical entry (content) needs to be structured consistent with the defined lexicon structure. For new lexica, the structure of a LEXUS lexicon is predefined and follows the LMF schema given in Figure 1.1 (a). Each of the elements in the schema can be expanded according to the wishes of the creator. Basically there are two ways of elaborating on such a model. You can either add a Data Component, which can be compared to a folder, or a Data Category, which can be compared to a sheet to store relevant linguistic information. The lexicon structure is an hierarchy, in which Data Components are higher in the structure than Data Categories. The former ones are used in order to group elements that have something in common. Form and Sense are the two most basic Data Components. To them you could link a number of other more specific Data Components, which in turn can include other sub-components or Data Categories. In the end all the lexical information of an entry is stored as values of Data Categories. It is up to the user to define the structure of the lexicon and in principle there are no restrictions on it whatsoever. In Figure 1.1 (b) we show a simple example of such a structure.

More information on LMF can be obtained from the LMF web site [<http://www.lexicalmarkupframework.org/>].

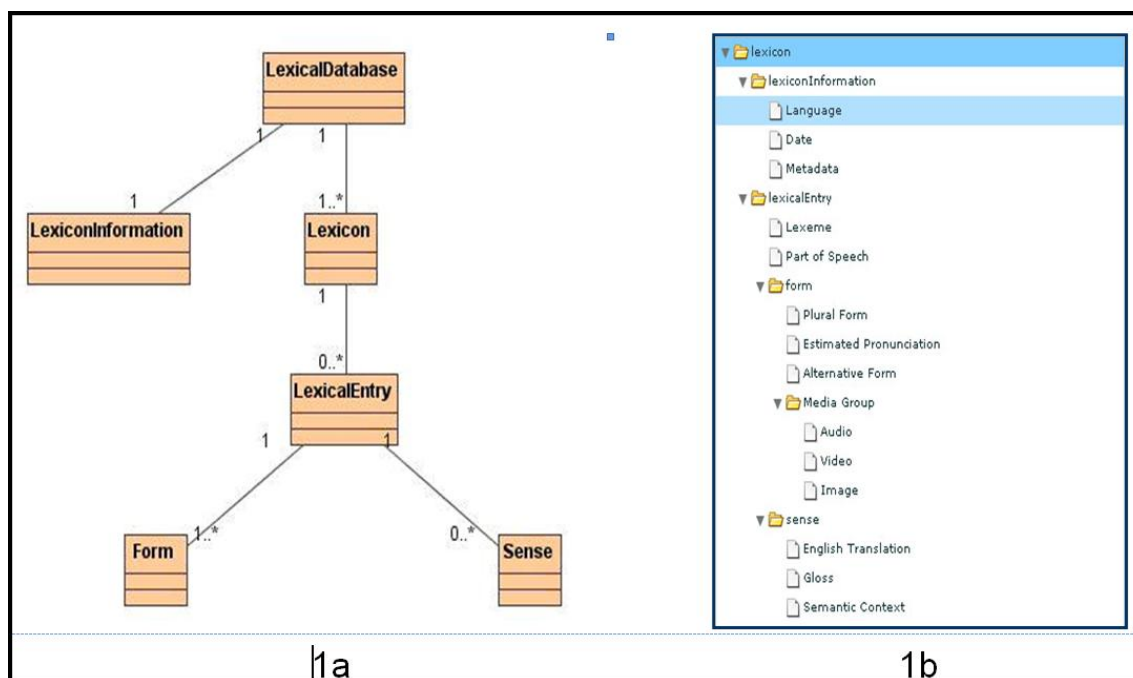


Figure 1.1. LMF lexicon schema (1a) and User defined LEXUS structure (1b)

Behind LEXUS is the tool ViCoS, visualizing conceptual spaces. With ViCoS you may extend your lexicon with a set of relations between lexical entries or between specific data category labels within the lexical entries. ViCoS thus allows the creation of a semantic network and the navigation through this network. Since ViCoS also allows the creation of user or lexicon specific relation types, the network that can be created go beyond the semantic into the cultural level. The created conceptual spaces show the words or attributes within the cultural contexts. For more information on ViCoS or the ViCoS manual see the LAT page: <http://www.lat-mpi.eu/tools/vicos>

For the best functioning of LEXUS and ViCoS, we advice to use the newest version of Mozilla Firefox browser with a minimum screen resolution of 1280 x 1024.

1.2. Registration and login

In LEXUS users need a workspace where the lexica are stored and where they can be shared with other users. To get this workspace, users have to register first. For now, the registration is not automated. Send an e-mail to Alexander Koenig [<http://www.mpi.nl/people/koenig-alexander>] from the Technical Group of the MPI, in order to get a workspace with a user name and pw.

1.3. LEXUS workspace

Once you have logged in, you will be directed to your LEXUS workspace. In the following chapters, we will show how you can (1) Create new lexica (See Section 2.1) (2) Add and modify lexical entries (See Section 2.3.2) (3) Import lexica from Toolbox or XML (See Chapter 3) formatted lexica, (4) Search through one or multiple lexica (See Chapter 4) and create Filters (See Section 4.2) on word lists (5) Create and apply sort orders (See Chapter 5) and (6) set Access rights (See Chapter 6) to other registered LEXUS users.

Before that, let us have a brief look at the workspace itself to see how it is organized:

The workspace opens by default with the left frame in the **Lexica** tab (Figure 1.2), in which all the lexica available to you are listed in alphabetical order. The lexica can include lexica from other workspaces for which you have read and/or edit rights. These so called 'shared' lexica are indicated with an icon of a hand holding a blue book. Lexica for which you are the first owner are indicated with a blue book icon.

When one of the lexica is highlighted, the **LEXICON** tab in the right frame displays information like Name, Description and Notes). Also in the right frame you will find the tabs **Readers** and **Writers**, which is where the access rights for other LEXUS users can be defined for the selected lexicon (See Chapter 6).

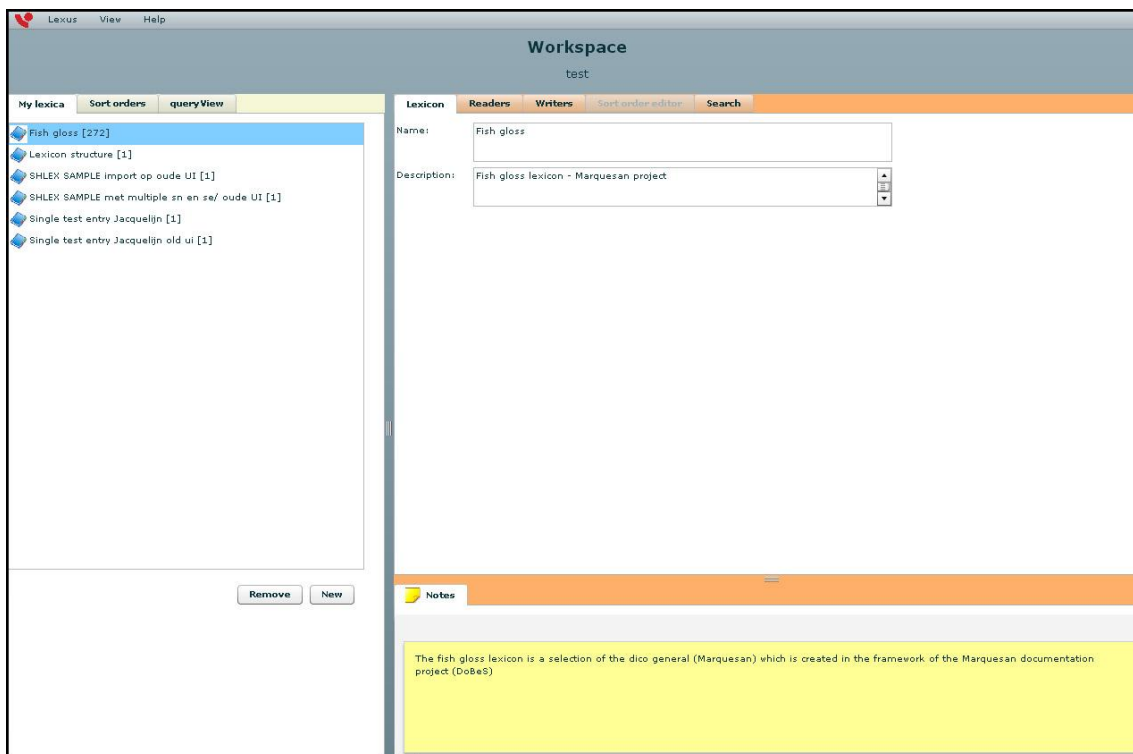


Figure 1.2. LEXUS workspace

The second tab of the left frame is **Sort orders**. It shows the sort orders that have been created within the workspace. The **Sort order editor** becomes active when in the **Sort orders** tab is selected: using this editor you may modify existing sort orders or create new ones. The sort orders can be used in all lexica in the workspace (See Section 5.2).

The third tab in the left frame is **Queries and searching**. In the window under the tab you may specify queries over one or multiple lexica (See Section 4.1). Queries are predefined searches which can be stored and re-used. **Ad hoc search** can be either used to do a simple search using keywords (See Section 4.1).

In top frame of the workspace, you will find the LEXUS main menu: **File**, **Switch to**, and **Help**. **File** is rather straight forward, where you will find **Save** and **Import** (see Section 3.1 and Section 3.2). **Switch to** allows you to switch between different editors in LEXUS and the workspace, and to access ViCoS (See Chapter 7). **Help** redirects to the online manual. To the right on the main menu bar you see the workspace username, from where you may **Log out**. The LEXUS icon holds information on the LEXUS version. Please check the version number with the version number of the manual you are using.

1.4. Workflow

In LEXUS some functionalities are dependent on one another. Therefore they will not work properly until some other operation has been performed. For this reason, it is important to follow a general workflow when working with LEXUS:

1. Create new lexicon (or import a lexicon)
2. Define the structure of the lexicon (or make sure that the structure of the imported lexicon has been kept)
3. Define the list view

4. Define the lexical entry view
5. Add new entries and media
6. Define and apply the sort orders
7. Other functionalities

Chapter 2. Create lexica

2.1. New lexica

New lexica are created in the LEXUS workspace. In the left frame Lexica tab click the + icon. This will open a new window, where you need to specify the name of the new lexicon, may give a description to it and add any other necessary comments in the Note field (Figure 2.1). Once you have clicked Save the lexicon will appear in the My Lexica list. The icon in front of the lexicon indicates that it is a new lexicon: no structure or content has been defined yet.

In case you would like to remove any of the lexica from your workspace, you can select it in the list in the left frame Lexica tab and click the - icon. If you remove a large lexicon, it can take a while until the lexicon is really removed. While LEXUS is in the process of removing the lexicon, the lexicon still appears in the lexicon list in the workspace, but with a different icon (the blue book held by a hand).

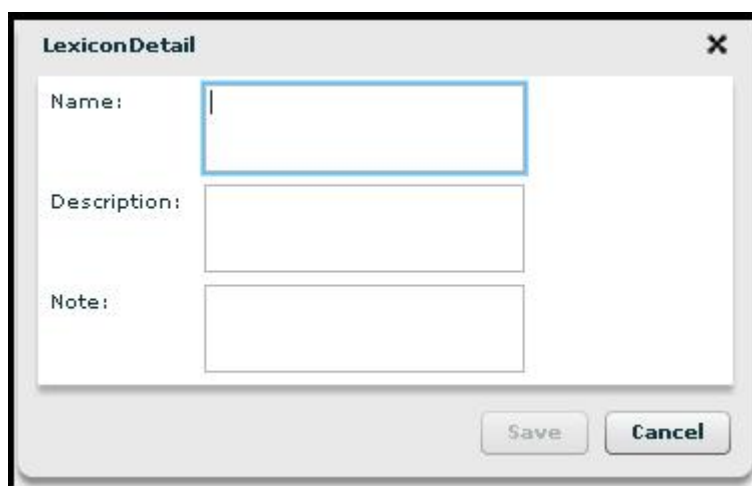


Figure 2.1. New lexicon information panel

2.2. Defining a lexicon structure

In LEXUS a lexicon consists of a structure and of content. You will always need to define a structure for your lexicon before you can add content. The default structure of a new lexicon is a basic LMF structure. A lexicon consists of two data components: LexiconInformation and LexicalEntry. LexicalEntry in its turn consists of two data components: Form and Sense. Data components are holders for data categories, which are the actual values holders. Compare this to a folder and file system. The data components are the folders and the datacategories are the files, which hold the content.

Its up to you to decide what you would want the structure of your lexicon to be. Group data categories which relate to the form of the lexical entry (like orthography, plural, but even image etc) under Form and group data categories related to meaning under Sense (e.g. gloss, definition). For first time users with LMF we refer to the LMF web site [<http://www.lexicalmarkupframework.org/>]. If you wish LEXUS also allows you to completely ignore the LMF framework. Just leave the Form and Sense component for what they are and build your structure directly under LexicalEntry.

2.2.1. Schema and View Editor

To define the structure of the lexicon, select the new lexicon in the lexicon list and switch to the Schema and View Editor (in the Switch to menu in the top frame). The editor window has two frames. The left frame

shows the default lexicon structure, consisting of a Form and Sense component. The right frame displays the information on the selected element of the schema. The information contains the name of the element, a description and specifies whether the element is mandatory in the lexical entries of the lexicon.

2.2.2. New elements

The default structure needs to be extended according to your requirements. First select from the default schema the element (Form, Sense etc.) under which the new item should be created and click the + icon below the tree. In the pop-up, select whether the new element will be a Data Component or a Data Category, give it a Name and a Description (see Figure 2.2).

Figure 2.2. New schema element

2.2.3. Modify an existing structure element and use the data category repository

To modify an existing item in the structure, select the element in the structure in the left frame, the right frame opens in the Schema Element tab. You can modify all the previously added information about an item in the structure.

In case of Data Categories you may decide to change from the user-defined type of Data Category to the predefined Data Categories from the Toolbox MDF categories or the ISO 12620 registry. Changing to a Toolbox defined data category is useful in those cases where you need maximum interoperability between the newly created lexicon and lexica in your workspace which were imported from the Toolbox application. Changing to an ISO 12620 data category enhances interoperability and standardizes your lexicon to a wider accepted conceptual format.

To change to a Toolbox data category, select the element in the structure in the left frame. In the right frame, under Type click 'Toolbox'. Select the Toolbox data category from the list. Note that the description field will fill with information on the Toolbox element.

To change to an ISO 12620 element, select the element in the structure in the left frame. In the right frame, under **Type** click 'ISO12620'. Fill in the name of the data category in the search box, select in which repositories you want to search and select one of the elements from the search results (see Figure 2.3). In this example we replace the user created data category 'head word' by the ISO cat data category 'word'. A description of the data category is given in the search result list. For more detailed information on the data category a link to the ISOcat web site is given under the magnifier glass right of the search results.

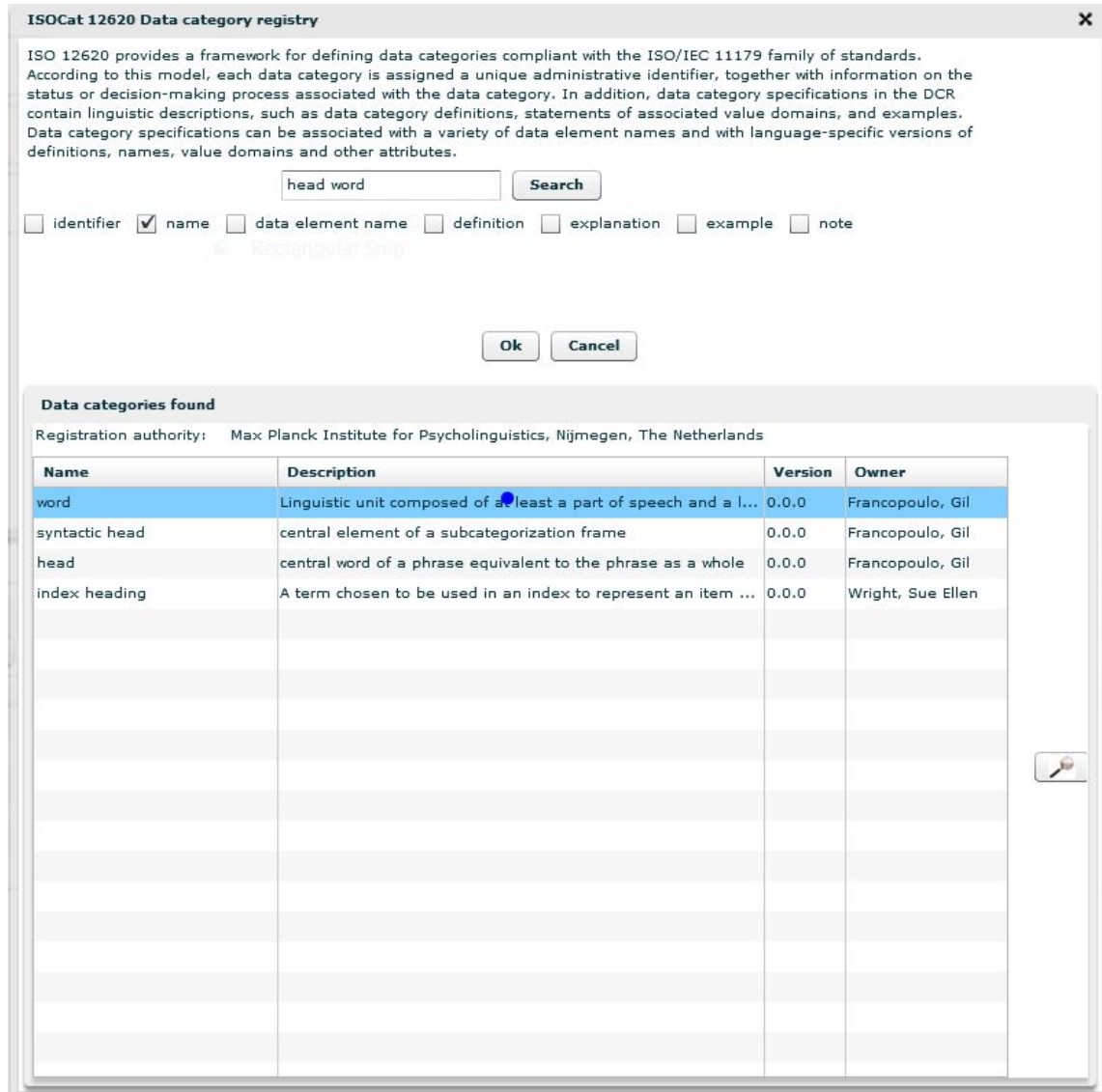


Figure 2.3. ISO Data category selection

2.2.4. Defining Word List View

In LEXUS the lexical entries are displayed in a word list. Once the structure of the lexicon has been created (and before you start to add lexical entries), we advise you to create a (preliminary) view of the items in the word list. You can define this view in right frame of the **Schema and View Editor**, select the **List View** tab. First you should decide which data categories from the structure of the lexicon should be used to represent the lexical entries in the word list. Usually the 'head word' data category is used for this. Drag the data category from the tree into the Preview window. When the Data Category is dropped and still highlighted in the structure on the left, you can change the properties of the display: font size, font type, color, bold, italic etc (See Figure 2.4)

In the same manner you can add more data categories to the **List view** and change their layout. Beware not to add too many data categories to the view, since this will create a chaotic view in the word list. Also

be aware that the order in which you add data categories to the list view should follow the structure of the lexicon. E.g. If your structure consists of a lexeme with a sub structure, followed by a sub-entry with a sub structure and you wish to display the lexeme with its gloss and the sub entry with its gloss, insert them into the List View Preview window in the order: Lexeme, gloss, Sub Entry, sub entry gloss (a bit of try and error will give the best result).

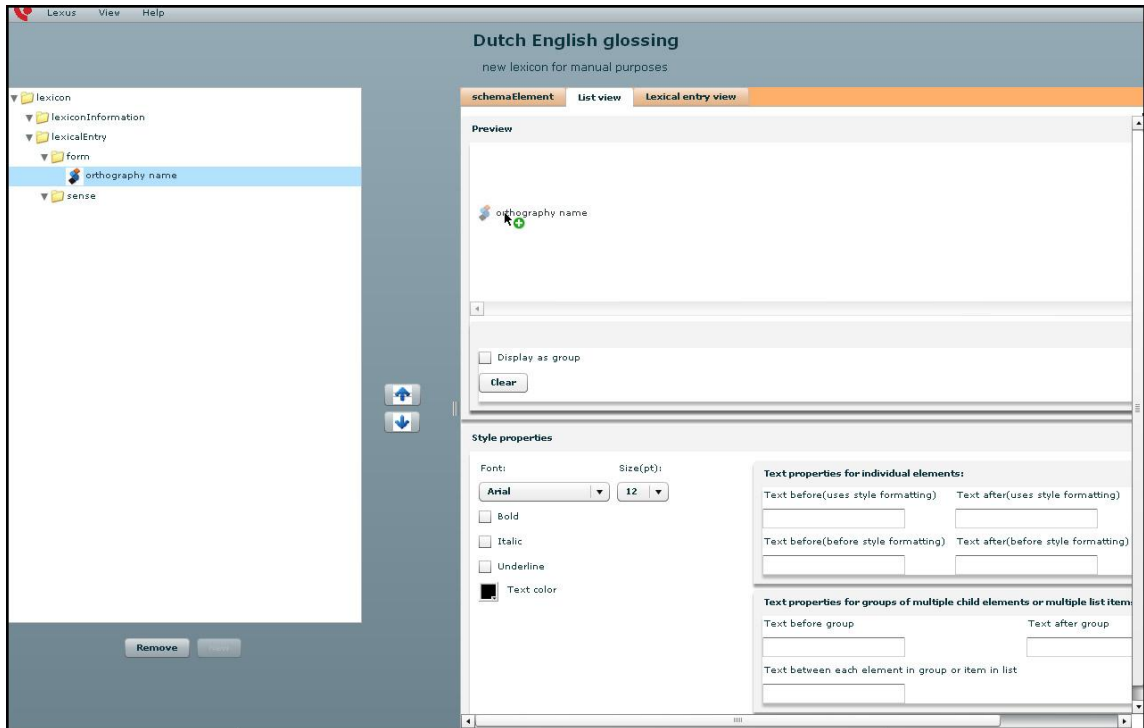


Figure 2.4. Defining the List View

Apart from showing the values of the data categories in the list, users can add additional text properties. For instance you might want to put a particular bit of information into brackets. For that the Text Properties windows can be used: Select the data category, which you have added previously to the List View and enter the required characters in the boxes. Users can add characters before and after the data category. Moreover, they can choose from two options: the characters can either use the formatting of the Data Category or keep their original style. In the example in Figure 2.5 the round brackets were added in such a way that they have the same color/style as the data category Lexeme, while the dashes keep their own formatting.

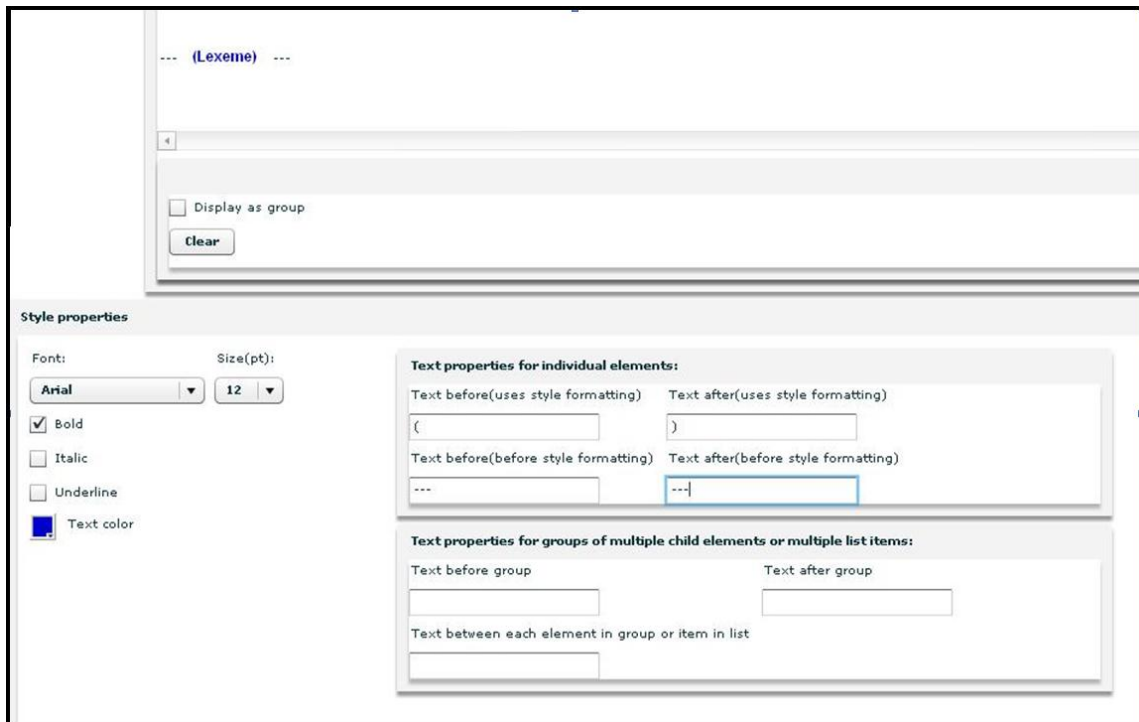


Figure 2.5. Adding free text to the word list view

It is possible to indicate in the wordlist that there is multi-media attached to the lexical entry. To do this, you need to select the element to which you attach the multi-media element for the word list view. Beware that the wordlist view will then also show the text of the element.

2.2.5. Defining the Lexical Entry View

After you have defined the list view, you also need to define the so called Lexical entry view. The Lexical entry view is the representation of the lexical entry on the screen, after it has been selected from the word list. This means it is a far more elaborate view of the lexical entry.

There are two ways of defining the lexical entry view. The users can either use an analogical editor as for the word list view. To do that you have to go to the **Schema and view editor** and choose the **Lexical entry view** tab in the right frame. Here, however, the data categories are not dragged from the structure tree to the Preview window, but chosen from the **Insert Data** category box at the bottom of the Preview window. Their properties can be changed using the menu bar on top of the page (see Figure 2.6)

The preferred way of defining lexical entry views however, is to create lay-outs in an external html editor (like e.g. Dreamweaver or NVU) and import them into LEXUS. The creation of such lexical entry views gives much better results, but is more complicated than using the editor. The Technical Group of the MPI provides support for this. We can give you example html files with views, or help you to create the best view for your lexicon. Send an e-mail to Alexander Koenig [<http://www.mpi.nl/people/koenig-alexander>] for help.

Once you have an html file, its easy to import the view into LEXUS. Go to the **Schema and View Editor** and choose the **Lexical entry view** tab in the right frame. Check the box **From local file system** and browse to the html view.

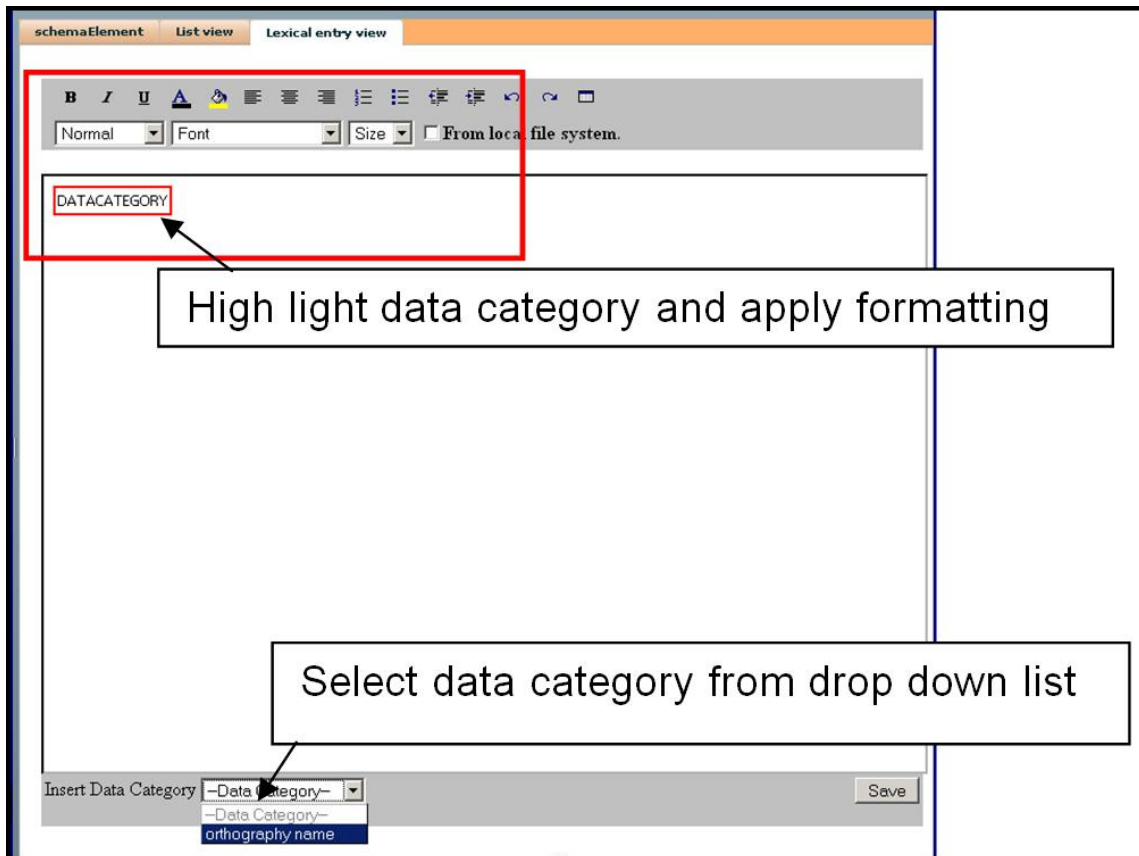


Figure 2.6. Lexical entry view editor

2.3. Lexical entries

After the structure of the lexicon has been created (see Section 2.2) and after the initial views have been defined (see Section 2.2.4) its now time to start to add the lexical entries. To do this, select the lexicon in the list in the Workspace and switch to the Lexicon editor using the Switch to menu in the top frame.

2.3.1. Lexicon editor

The Lexicon editor window has a two frames (see Figure 2.7). In the left frame there are two tabs: Lexicon and Filters.

In the Lexicon tab is the (initially empty) word list. Above the word list there are four options that you may choose from: (1) You can choose to display the word list of a different lexicon by using the drop down list in the box <switch lexicon>, (2) You can choose a selection of the words, based on initial character by selecting the character from the drop down list <grapheme>, (3) You can display a selection of the lexical entries by selecting a Filter (Section 4.2)or finally (4) you may switch to the structure view of the selected lexical entry (as opposed to the list view).

In the right frame there are two tabs: Lexical Entry View and Lexical entry. The first one displays the lexical entry in a configuration as defined in the Lexical Entry View editor (see Section 2.2.5). The second one displays the values of each data category of the lexical entry in a structured manner.



Figure 2.7. Lexicon editor, frames and options

2.3.2. Adding new lexical entries

To add a new lexical entry, click on the + icon under the word list. A new entry will appear in the list, with value 'unspecified'. In the right frame the Lexical entry tab gives you the option to add values for the data categories in the boxes (see Figure 2.8).

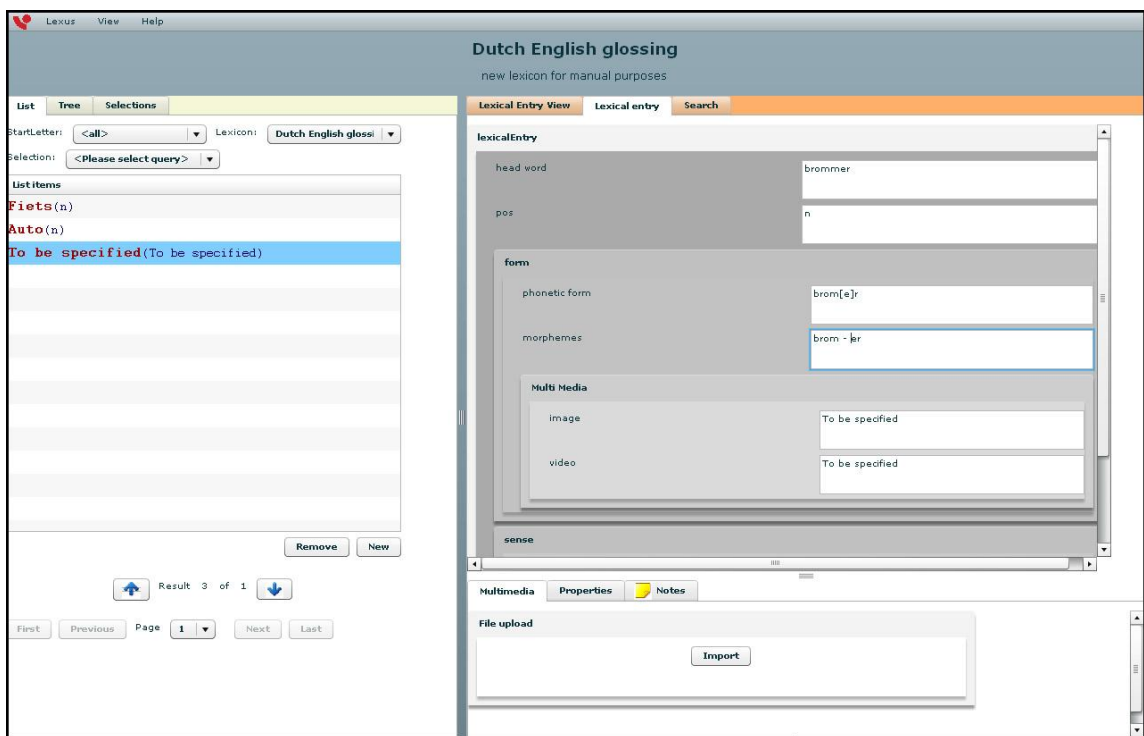


Figure 2.8. New lexical entry (brommer)

2.3.3. Modifying lexical entries

To change the values of existing lexical entries is similar to adding new lexical entries. In the word list, select the entry that you wish to edit, in the right frame open the Lexical Entry tab and adjust the values. To save the changes, select File --> Save from the menu in the top frame.

Besides changing the values of the lexical entries, you might also want to add data categories or components. Now remember that the structure of the lexical entry always needs to be consistent with the structure defined for the lexicon (see Section 2.2), but some elements can appear more than once in one lexical entry, for instance a lexical entry can have more than one sense.

In the word list select the lexical entry and switch to 'structure view'. Select the data component to which you want to add a new element. At the bottom of the frame is an option to add features. Add the required element and Save from File and Save in the top frame (see Figure 2.9).

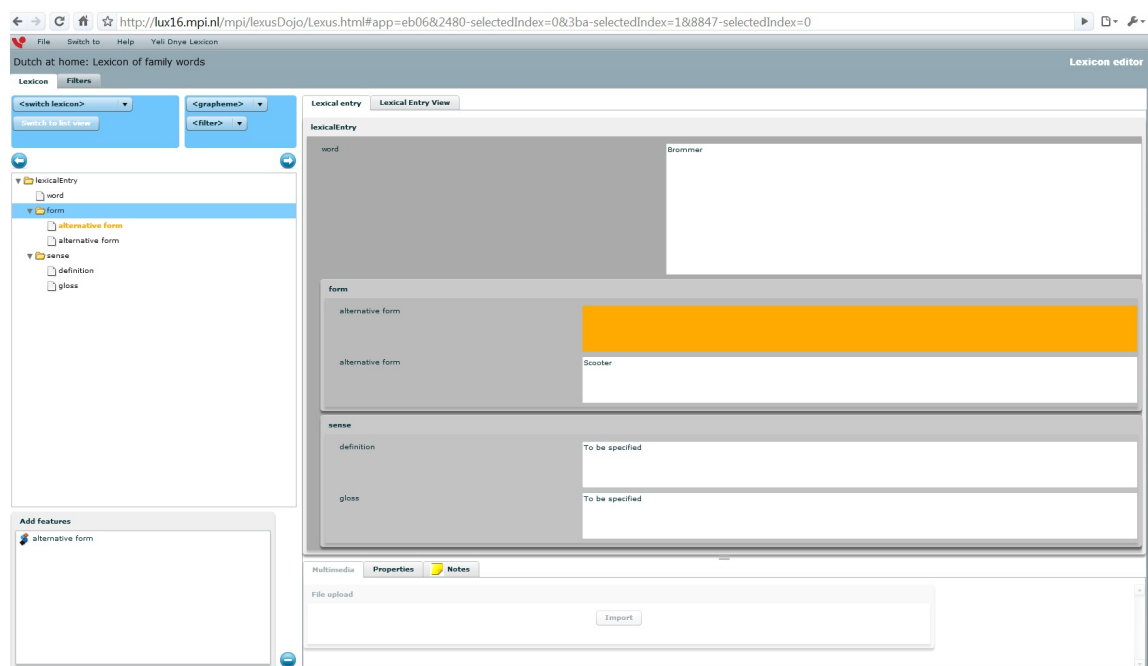


Figure 2.9. Modifying the structure of a lexical entry

2.4. Adding multi media to lexical entries

In LEXUS you may add three types of multi media: images, sound files and video files. It is wise to create a separate data category for the multi media in the lexicon structure. An example of this is given in Figure 2.10, where a Multi Media component with two data categories: image and video, is linked to Form. How to add data categories and components to your structure is explained in in Section 2.2.

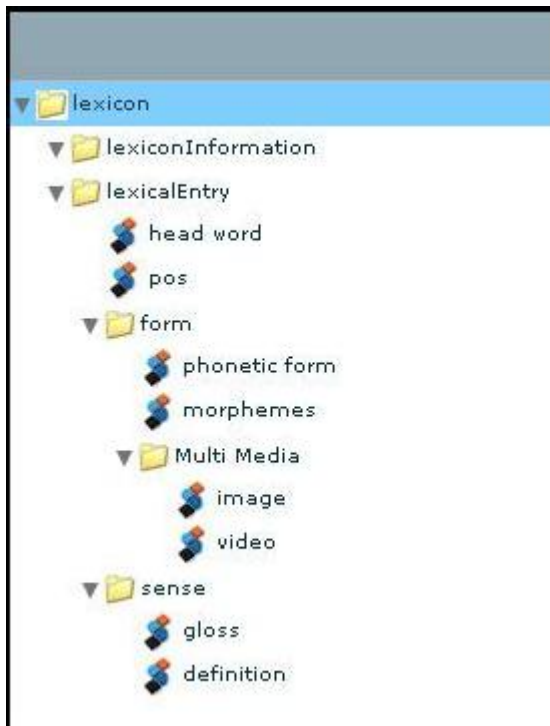


Figure 2.10. Multi media in the lexicon structure

To add multi media to the lexical entry, go to the Lexicon editor and select the lexical entry from the word list. In the right frame select the **Lexical Entry** tab and move to the data category where you want to add the multi media. In the box enter a header for the media. Next, in the lower part of the right frame, select the **Multimedia** tab and click import. LEXUS prompts you to choose between uploading the media from a local resource into the LEXUS data base or create a link to media stored in the MPI archive for linguistic resources.

(1) From a local resource: click **select resource** and browse to the image, sound file or video file. After uploading save in the top frame menu: **File --> Save** If you have defined the lexical entry view such that it will display the value of the media data category, it will show the media in a boxed frame which includes the header (see Figure 2.11). The image is not scalable, but has a standard width.

Note that since LEXUS allows to add multi media to every data category in the structure, you do not **NEED** to create a separate data category for the multimedia. You could also add the image to e.g. the head word data category. In the lexical entry view this would result in a box around the headword and the media file.

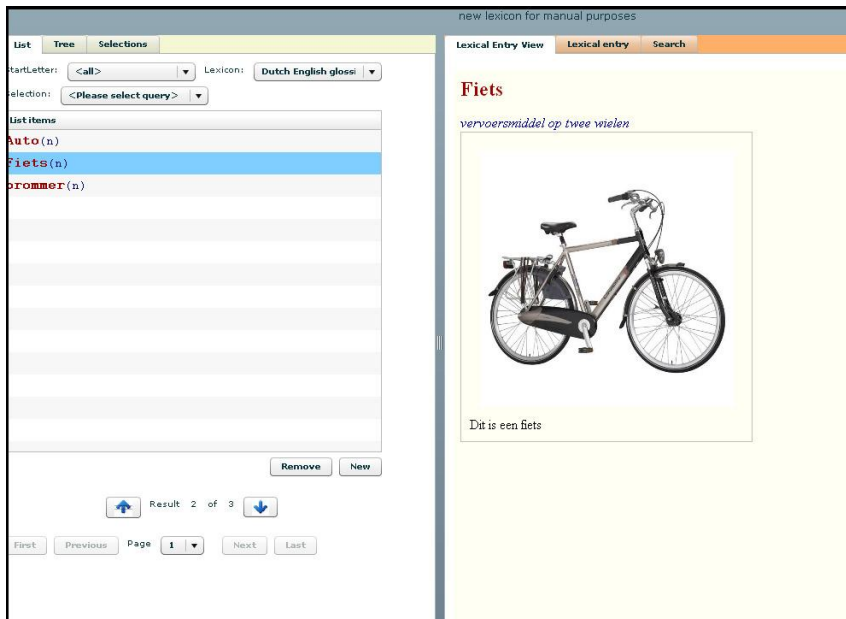


Figure 2.11. Image in the lexical entry view

(2) Linking to the archive: Select Archive in the Import Resource pop-up. LEXUS prompts you for the internal node ID in the MPI archive. To find this node ID, open a new tab in your internet browser and open the MPI archive at <http://corpus1.mpi.nl> in the corpus tree browse to the node where you want to link to, e.g. an ELAN annotation file. Click on the resource in the tree, and copy from the right frame, the value of the internal nodeid (e.g. MPI600401#). Copy this ID in the LEXUS prompt and select: get resource. When you have selected an ELAN file, LEXUS will also prompt you to feed the start time and end time of the fragment you wish to link to LEXUS. When you ignore this LEXUS will jump to the beginning of the file. After uploading save in the top frame menu: File --> Save

Chapter 3. Importing and exporting lexica from other formats

3.1. Importing lexica from Toolbox

Apart from creating a new lexicon, LEXUS also gives the option to import already existing lexica from two formats: Toolbox and XML. Importing from Toolbox requires that the lexical entries in the Toolbox database are consistent with the structure defined in the Toolbox *.typ file. Since Toolbox allows 'inconsistent' lexical entries, it is really important the lexicon is checked before importing. There is a separate manual on how to do this. Please check the LEXUS website (www.mpi.nl/lat/tools/lexus [<http://www.mpi.nl/lat/tools/lexus>]).

Assuming that the lexicon is consistent with the *.typ file: choose File and Import Toolbox from the menu in the top frame. A window will pop up, in which you can first browse for the Type File and next for the Data File (Figure 3.1).

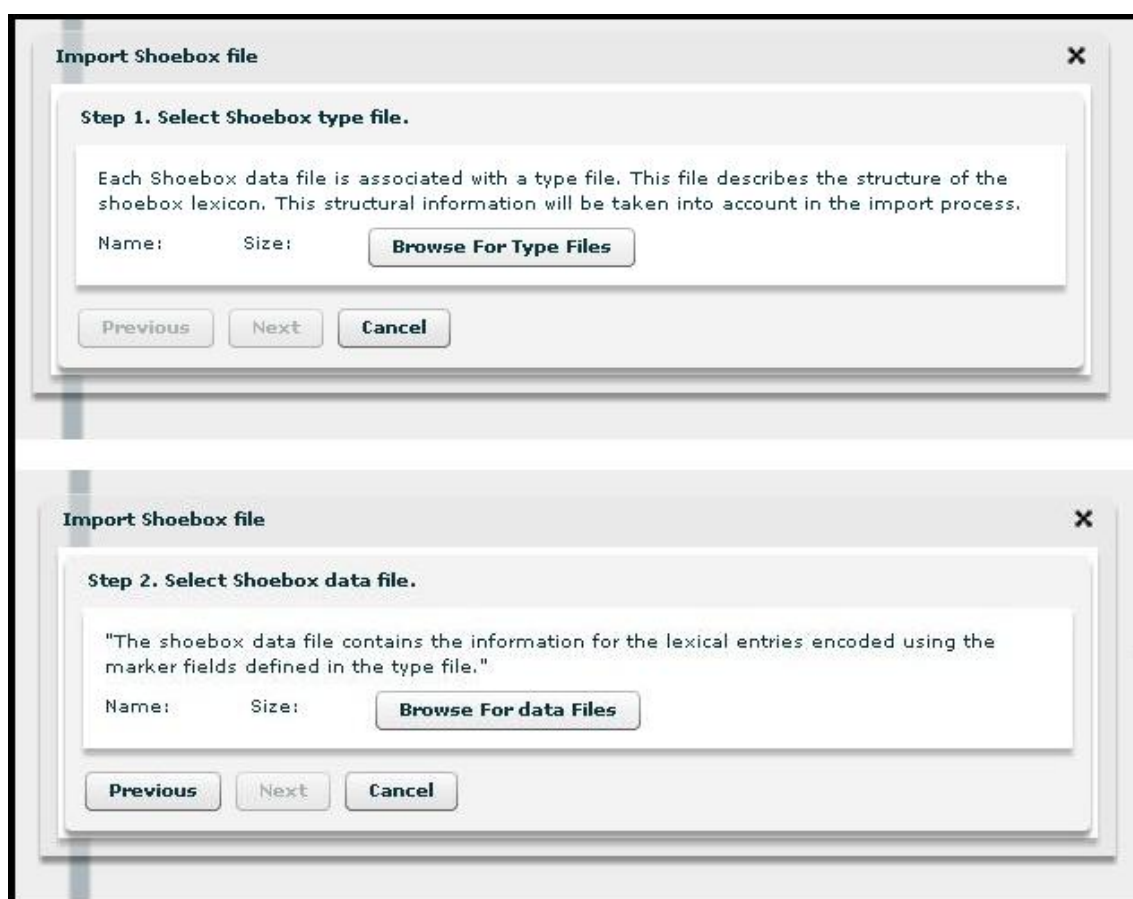


Figure 3.1. Toolbox import pop-ups

After that you have the option to make the structure compatible with the LMF format. For Toolbox lexica this is often particularly difficult as the Toolbox structure tends to mix the Form and Sense categories of lexical items and separating them into the default LMF structure is often not possible. In these cases, we therefore strongly advise to choose all the markers from your structure and link them directly under Lexical Entry. You can choose all the markers at once by clicking on the first one and then clicking on the last one in the list while pressing Shift. Then simply drag them and drop them on the Lexical Entry Data Component (Figure 3.2).

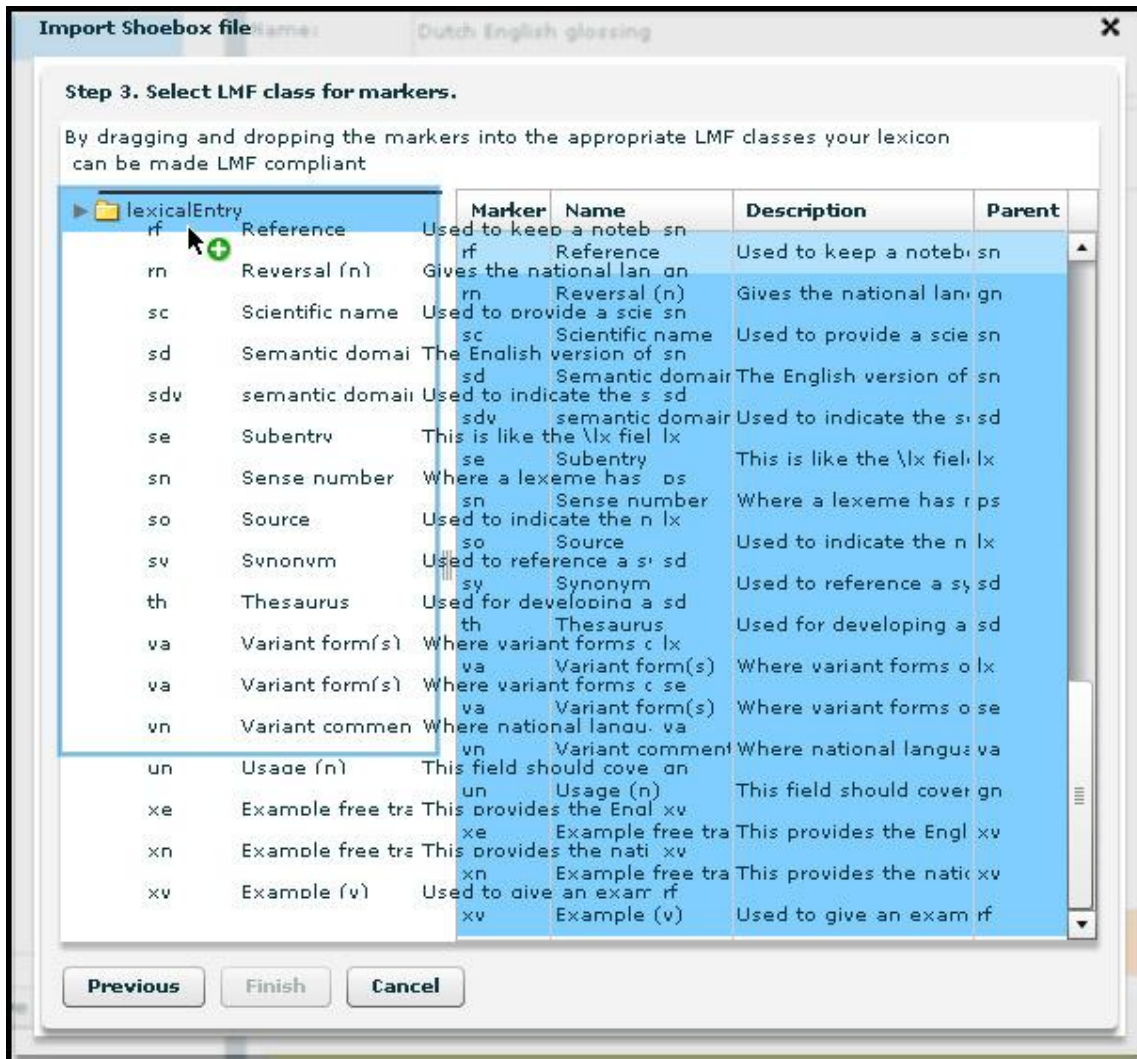


Figure 3.2. Toolbox import, all elements under Lexical Entry

Click Finish. The import process can take from 2 minutes up to an hour, depending on the size of the lexicon. While the import process takes place, the lexicon will appear in the lexicon list in the *Workspace*. The number behind the lexicon name indicates the total number of lexical items already imported. Once all the data have been successfully imported, you can define the word list view and lexical entry view.

The structure of the imported lexicon is based on the structure defined in the Toolbox typ file. However, in the hierarchy, there is a number of elements that are on a par - they are sister nodes linked under one mother node. For LEXUS they are equal, therefore they are ordered randomly. The users can however define the layout of the structure in LEXUS: that is decide in what order the data categories and data components of the same level should appear. It is for instance quite natural that the users will want the ps group under lexeme and the ps group under subentry to look the same. In order to do that, in the *Workspace* select lexicon from the list and choose the *Schema* and *view editor* from the main menu *Switch to*. The structure will be visible in the left frame. Select the data component or data category you wish to move and use the up and down arrow buttons next to the structure to move the element. You may also delete elements from the structure: select the data category or component in the tree and click on the - button below the tree. These functions are shown in Figure 3.3.

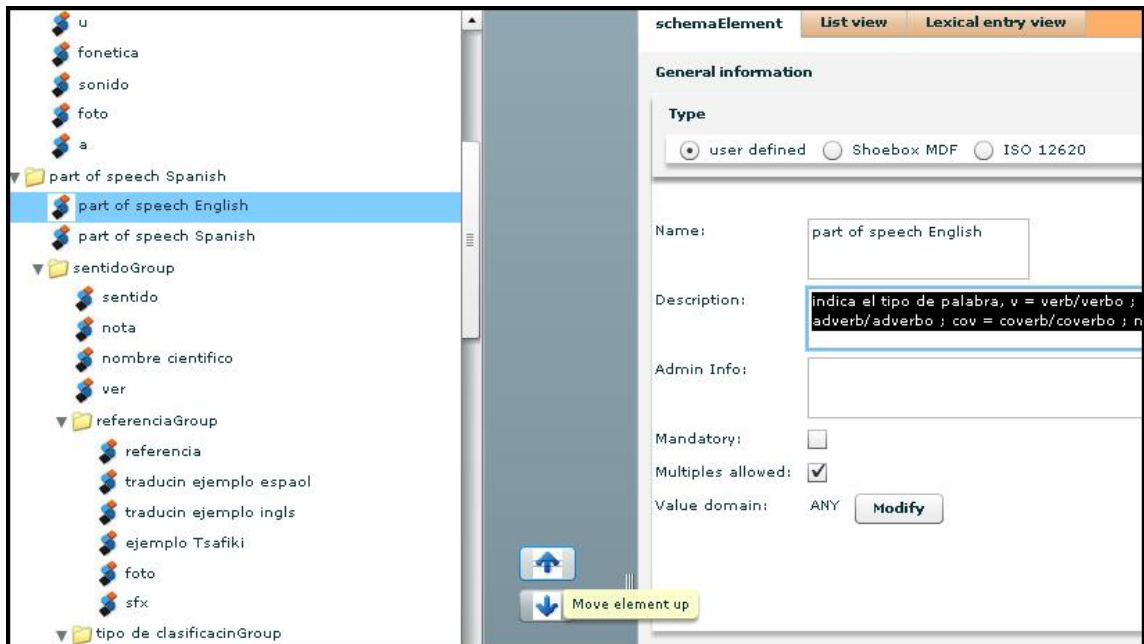


Figure 3.3. Changing the layout of the LEXUS structure

3.2. Importing lexica from XML

In order to import an XML you require both an XML data file, and an XSD schema file for which the data file validates. If you have both, select Import XML from File in the main menu. First you are prompted to browse to the XML data file, followed by a prompt for the XSD schema file. After that you have the option to make the structure compatible with the LMF format. You can drag the elements from their original position in the XML structure into their new position in the LMF structure. Figure 3.4. In the import frame complex XML elements are displayed in blue circles, simple XML elements are green circles. In the imported structure, the complex elements will become Data Components and the simple elements will become data categories.

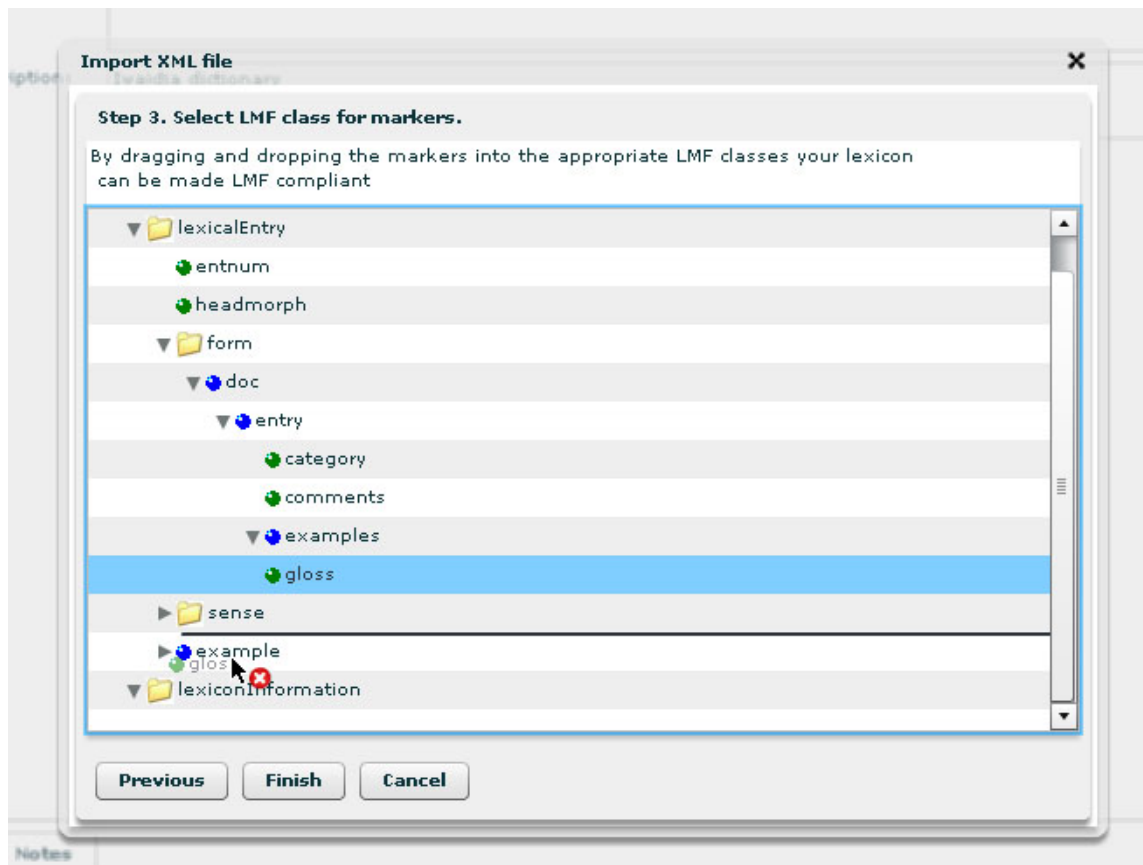


Figure 3.4. XML import: restructuring the elements

3.3. Exporting lexica to XML

Lexica can be exported to XML. To do this: open the lexicon for which you want to create an XML export file in the Lexicon Editor. Next select XML export under File in the main menu. A dialogue frame opens which asks you to specify the XML export file name and the folder where to store the export file. You will get a DOWNLOAD COMPLETE message when the download is ready. Use an external XML editor to create an XML schema file for the lexicon.

3.4. Import Toolbox data to an existing lexicon

You can import additional Toolbox data into an existing LEXUS lexicon. Do do this some precaution needs to taken and always run a test with one or two entries before you import a larger data set. The import cannot be undone easily. The only way to undo the import is by deleting the imported entries one by one. To import Toolbox data to an existing lexicon you need a Toolbox typ file and the Toolbox data file.

- (1) Select the lexicon in the Workspace and switch to the Lexicon Editor, where you can see the wordlist and lexical entry tabs.
- (2) Select File and Import Toolbox. In the Pop-up you are prompted to select the typ file and after that you are prompted to select the data file.
- (3) The next window is similar to (Figure 3.2). The Difference is that the import process recognizes the elements from the typ file and matches these to the elements in the existing LEXUS structure. The recognized elements have a green color, and the matching elements in the LEXUS structure have the marker name in

brackets behind the data category name (e.g. Lexeme [lx]). (Figure 3.5). You should check this for all the markers!

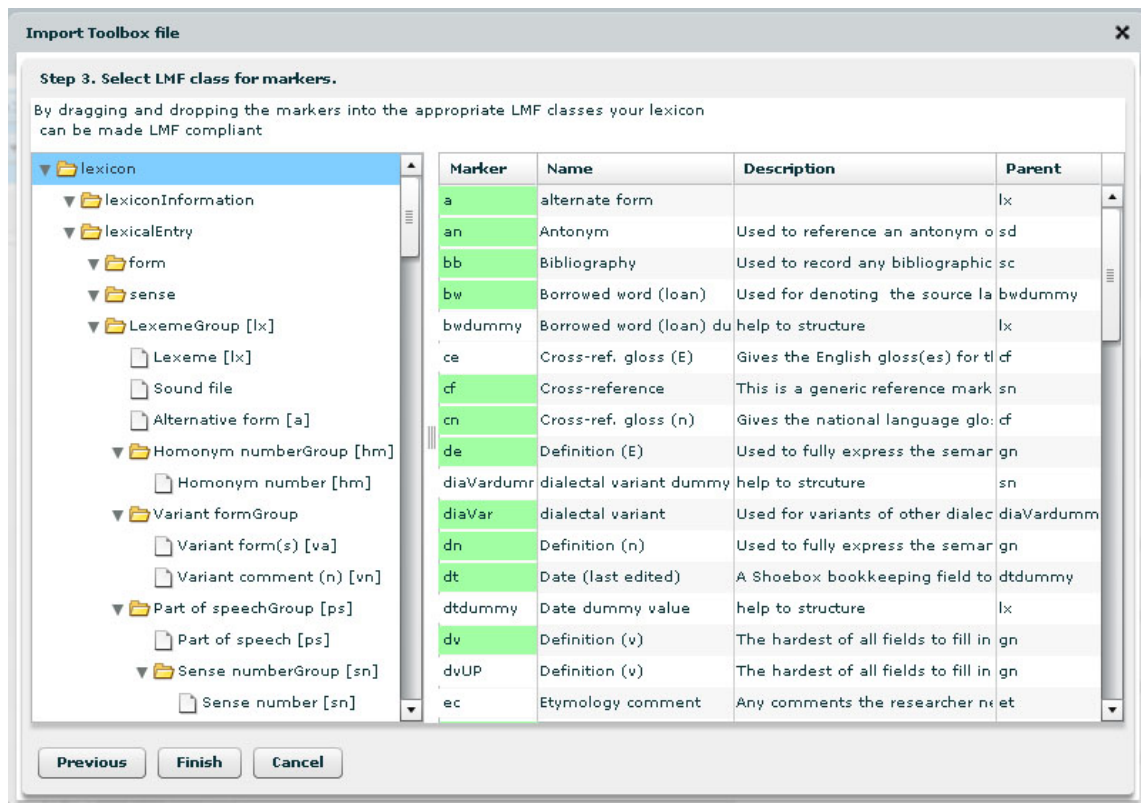


Figure 3.5. Toolbox data import to an existing lexicon

(4) There are two things that might not be correct, and should be corrected.

(4a) The LEXUS element might not have the marker in square brackets behind the data category name. If this is the case, abandon the import process (click cancel)

From the main menu, select **Switch to** and **Schema** and **View editor**. In the tree, select the element with the missing marker indication (e.g. lexeme). In the right frame change the **Type** of the marker from **User defined** to **Toolbox**. In the list with **Toolbox** markers that pops-up, select the correct marker (e.g. lexeme). Select **File** and **Save**. This should solve the problem of the missing marker. Go back to the **Toolbox** import to existing lexicon and start from the beginning.

(4b) The marker is not green in the **Toolbox** import window. In this case you should not abandon the import, but drag the non-green element to the **LEXUS** element where you want this marker to be in the tree. (Figure 3.6).

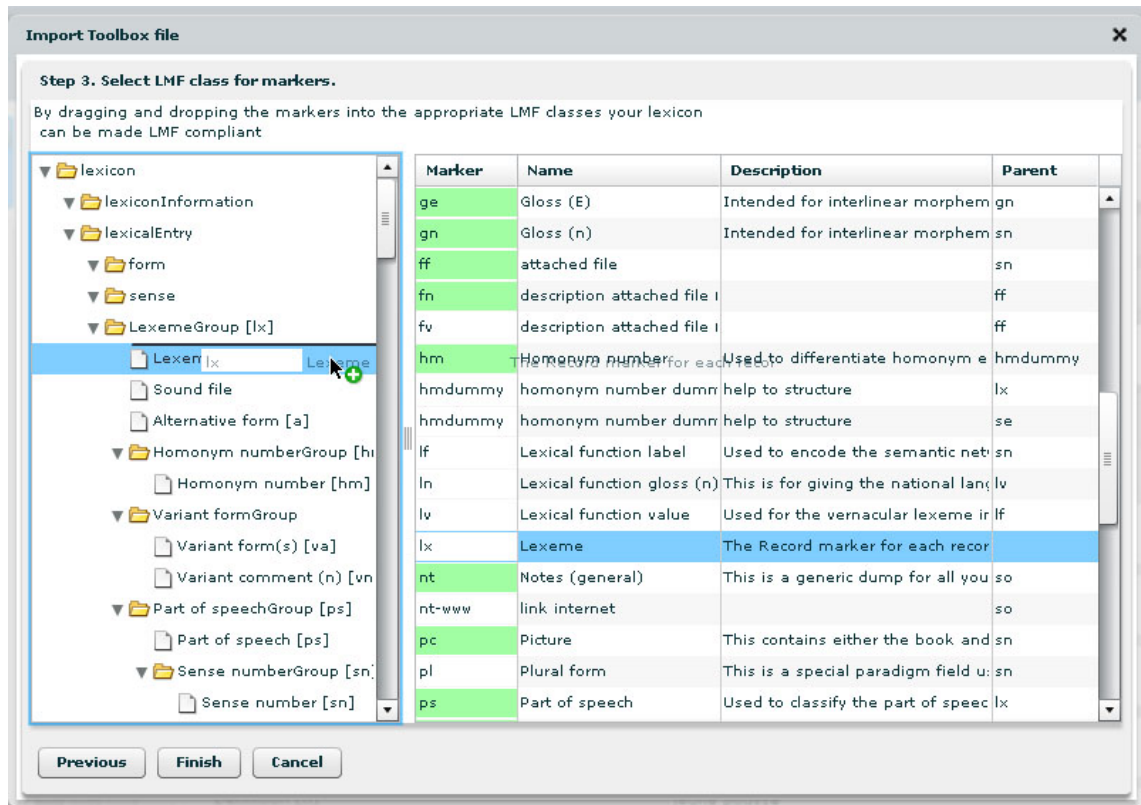


Figure 3.6. Drag and drop of 'non-green' marker in Import to existing lexicon process

(5) When you have made sure that both condition described under (4) are OK, you can click Finish. The import can take some time depending on the number of entries you are importing. In the Workspace lexicon list you can see when the import is done, when the number of lexical entries behind the lexicon is again one single number.

Chapter 4. Searching and filters

In LEXUS you can search and filter. Searching is at the workspace level, you can search either from a pre-defined query or from an adhoc search. Both are explained in Section 1 of this chapter. Filtering is at the lexicon level. You can define a filter on a word list, to display only a selection of the lexical entries. How to create filters and how to apply them is explained in Section 2 of this chapter.

4.1. Search

LEXUS provides search functionalities for searching all the lexica within your own workspace. This includes those lexica for which you have been granted rights to, but which originate from other workspaces (users). Search has two options: through an ad hoc (quick) search and through a query definer.

(1) Use Ad hoc search in case you want to do a relatively simple search in only one lexicon and in one data category. Ad hoc search is available in the Queries and searching tab of the Workspace.

(2) With the Query builder you may define more complex search queries, over multiple lexica and in more than one data category. The query can be saved and re-used. The query can be defined in the Queries and searching tab in the Workspace, and used for searches in the right frame in the Search with query tab.

4.1.1. Ad hoc search

In Ad hoc search, you need to specify: (1) the lexicon in which you want to search, (2) the data category, (3) the search value and (4) the type of search (is/contains/begins with/ends with). An example is shown in Figure 4.1.

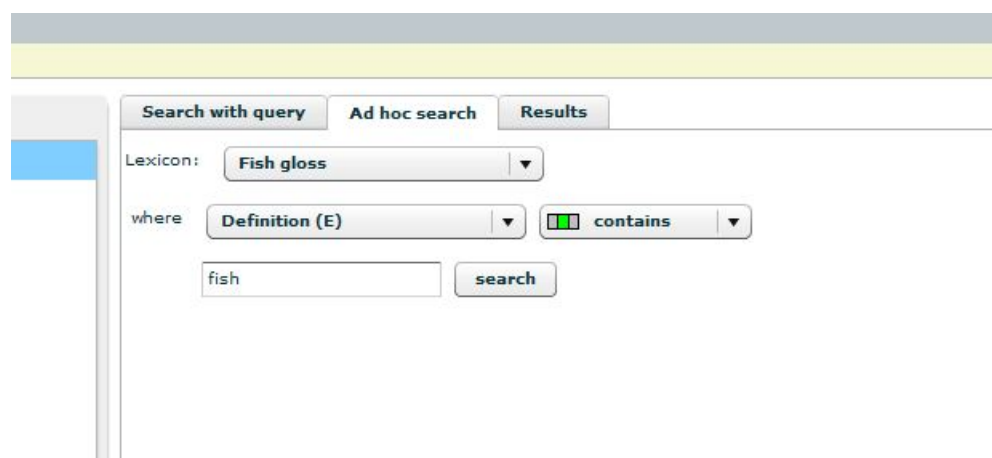


Figure 4.1. Ad hoc search

4.1.2. Creating queries

With the query builder you can create complicated searches on multiple lexica. The Query builder is still a bit buggy, but we try to explain exactly what to do, so if you follow this manual word by word, it should work. In the mean time we work on the bugs.

To create a query first go to Queries and searching in the left frame of the Workspace. In the left frame you will find a list of available queries in the workspace. The list is initially empty. Click + icon to create a new query. In the pop-up window, add a name for the query and if you want also add a description. E.g. if you want to create a query to select all lexical items which are defined as 'tree' in one or more of the lexica in your workspace, call the query 'Trees'. Click OK. Your new query now appears in the list. (Note

that when your list is not empty, and you create a new query by selecting the + icon, a different window pops-up. In this window: check the circle in front of Filter. The Name that will appear is the name of one of your previously created Queries, but simply overwrite this name (this is one of the buggy behaviors), and a new query will be created leaving the old one in tact.)

Now, your query is created, but not yet defined. Close all the pop-ups that have appeared (buggy again!). Select the new query in the list and click the + icon. Now select a lexicon in the drop downlist, for the first query constraint. Next select the data category in which the query needs to be performed, the value and the condition (not/is/contains etc.). You have now defined the first constraint of the query. If you want to define another constraint, you must first decide whether the second condition will be in an OR relation to the first (for example data category definition should contain 'fish' or 'net') or if you want the second constraint to be in an AND relation to the first ('fish' AND 'net').

In case you want an OR relation: Select the lexicon node of the query, and click the + icon. Select the data category for the second constraint, the value and the condition. Your query should not look like the Fish query in Figure 4.2. The Fish query in the image searched for entries in the lexicon Fish Gloss, that contain the value 'fish' in the data category 'Definition (E)' OR in the data category 'Definition (E) (se)'.

In case you want an AND relation: Select the data category node of the query and click the + icon. Select the data category for the second constraint, the value and the condition. The new condition will now be sub-ordinate to the first.

You can also search in multiple lexica, using an OR relation between them. To add a second query constraint from a second lexicon. Select the Query node in the tree and click the + icon. Select the lexicon for the second constraint from the drop down list and further follow the procedure as described above. The 'Tree query' shown in Figure 4.2 will search for: lexical entries in the lexicon 'Yeli Dnye demo lexicon' that contain the string 'tree' in the data category "description' OR lexical entries in the lexicon 'Coconut glossary' that contain the string 'tree' in the data category 'Definition(E)'.

Always save your created query from the File--> Save menu.

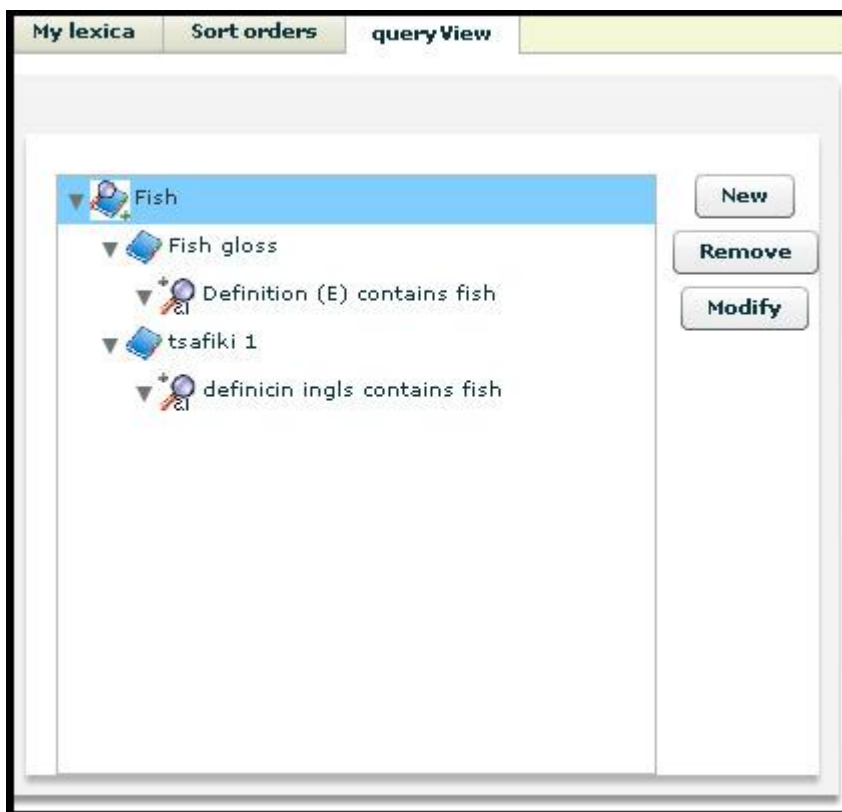


Figure 4.2. Query builder

4.1.3. Using queries

The queries are available for search in the **Search with query** tab of the right frame of **Queries and searching**. Select the query from the drop down list, a new tab **Results** will be added to the right frame. In the **Results** tab the search results are given in the form of one or more word lists. When you click on one of the items in the list, a new window will pop-up with a **details view** and a **Lexical entry view** of the Lexical Entry. Newly created queries will only be loaded into the drop down list after an reload of the system (usually F5 in your browser).

4.2. Filters

With LEXUS you may also create filters within lexica. For instance, you may create a filter which filters out all the lexical entries which have 'noun' as a value for the 'part of speech' data category. Applying this filter on the word list of a lexicon, will result in a word list of nouns only. Likewise you may create thematic filters, for instance to select only those objects from your word list which represent animals.

4.2.1. Creating filters

Since filters are specific for a certain lexicon, you need to define the filter in the **LEXICON** editor, under the tab **Filters** in the left frame. The process of creating a filter is similar to the process of creating a Query, with the sole exception that from the drop down list of lexica there is only one lexicon that you may select. Please check the Section **Creating queries**.

4.2.2. Using filters

The filters are available from the <filter> option just above the word list in the **Lexicon** editor. Select the filter from the drop down list and the word list will show the lexical entries that match the filter criteria. Like for the queries you need to reload the browser before you can use a newly created Filter.

Chapter 5. Characters and sorting

5.1. Using non-latin characters

The values of the data categories can be entered in non-latin characters.

For Windows users: first you need to add the language that you wish to use to the "Languages" in the "Regional and Languages options". To do this (in Windows XP): open the Control Panel and the Regional and Languages Options. On the Languages tab select Text Services and Input Languages and select Details. Under Installed Services click on Add and add the required language. Under Key Setting you can check how to switch between input languages. If you are using another version of Windows, or if the language you wish to use is not in the drop down list of available languages, you should check on the Windows Internet site on how to add languages to the Input Languages.

Once you have the language available in Windows, in LEXUS you need to select **Activate Unicode input** in the Menu option under **Help** in the top menu. The input boxes under the **Lexical Entry** tab of the **Lexical Entry Editor** should turn grey. We have not completely covered the switch issue, so sometimes the box stay white and you will not be able to enter non-latin characters. The work around is to change to the other tab, change back to the **Lexical Entry** tab and click the value bar of the data category. We are sorry for this inconvenience, but we thought it would be good to make the functionality available to you, while we try to repair the switch problem. Please give us feedback on the use of the Unicode input functionality.

5.2. Sort orders

In LEXUS you can create sort orders which are language specific. For every data category you can indicate which sort order should be applied. So the process of sorting in the right order has two steps: first create a sort order and then apply it to a data category.

5.3. Creating sort orders

Sort orders need to be created at the workspace level, so that they can be used on any data category in any of the lexica in the workspace. In the left frame of the **Workspace** select the **Sort orders** tab. In the list in the left frame you will see all the sort orders available for the workspace (initially this list is empty). To create a new sort order click the **+** icon and give a Name and Description to the sort order in the pop-up frame.

When you select the new sort order in the left fram, the right frame **Sort order editor** becomes active: the name and the description of the sort orders are heading a table which consists of two columns, each with a small character western alphabet. To adjust:

1. Remove characters from the list Select the character and click the **-** icon. The character will be removed from the list.
2. Define the characters to be sorted under a specific character. Select the initial character (e.g. a) in the left column. Define the list of characters that you would like to be sorted under the 'a'. e.g. AaÀàÁáÂâÃãÄä and add this list in the right column. Beware of the order in this list, LEXUS will prioritize (first in the list comes first in the sort order).
3. Using combined characters in the list When you are using e.g the combination of '-' and 'a' as in '-a' and you would like this to be sorted under 'a', then you should put this combination in square brackets in the list in the right column [-a]. A example of this is shown in Figure 5.1
4. Using combined characters as initial character In some languages a character exists only as a combination of two. E.g in Dutch there is a character [ij] combining the 'i' and the 'j'. Some people like to this as one character and sort it in the alphabet after the 'x' and before the 'y'. To do this move to the location in the alphabet where you want to add a character and click 'Add'. An empty line appears where you may add the

new character. In the left column there is no need to use the square brackets: the character will look like the example in the Figure. More complex sets of characters have been seen, especially in endangered languages. E.g the Iwaidja alphabet contains e.g. a combination of '^(\rt)a' which should be sorted under 'a'. Because the '(' symbol is used for development code, the combination needs to be formatted as '[^\(rt)a]' in the sort order table. For questions on special sort orders send an e-mail to Alexander Koenig [<http://www.mpi.nl/people/koenig-alexander>].

After you have created a sort order, make sure you Save it through the main menu in the top frame, File --> Save

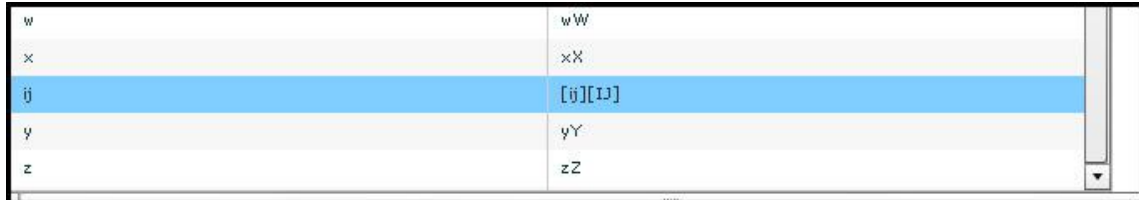


Figure 5.1. Setting sort order for combined characters

5.4. Applying sort orders

Sort orders are stored at the Workspace level; sort orders can be applied to the data category which is first defined in the wordlist view. After applying the sort order to this data category, the word list will be sorted according to the specified sort order. Before you can apply the sort order, you need to define the Word list view (see Section 2.2.4) . Next you may apply a sort order to the element defined. For the definition of the word list view you are already in the Schema and View editor, now go to the Schema element tab, and in the tree in the left frame, select the data category to which you want to apply a sort order. Once the data category is selected, in the right frame, bottom part, you can select a sort order. Save through File --> Save.

Chapter 6. Access rights

LEXUS allows you to share your lexica with other registered LEXUS users. You can either assign them read rights only, which will enable to user to view the lexicon in his or her workspace. You may also grant write rights to other users. This means that this user can open the lexicon in his or her workspace and make adjustments to it. There is no option to grant partial write right.

To grant rights for a lexicon, move to the **Workspace** and select the lexicon from the list in the left frame. In the right frame select the **Readers** tab (or the **Writers** tab in case you want to grant writing rights). In **Readers** you will see the currently assigned readers to the lexicon, which initially is just the workspace holder (you). Load users in **LEXUS users** to see all registered LEXUS users. Select the users that you wish to grant rights to and drag them to the **Readers** frame (see Figure 6.1).

In the lexicon list, lexica that have their origin in an external workspace have a different icon than the ones created in the workspace. They are indicated with a hand holding a blue book.

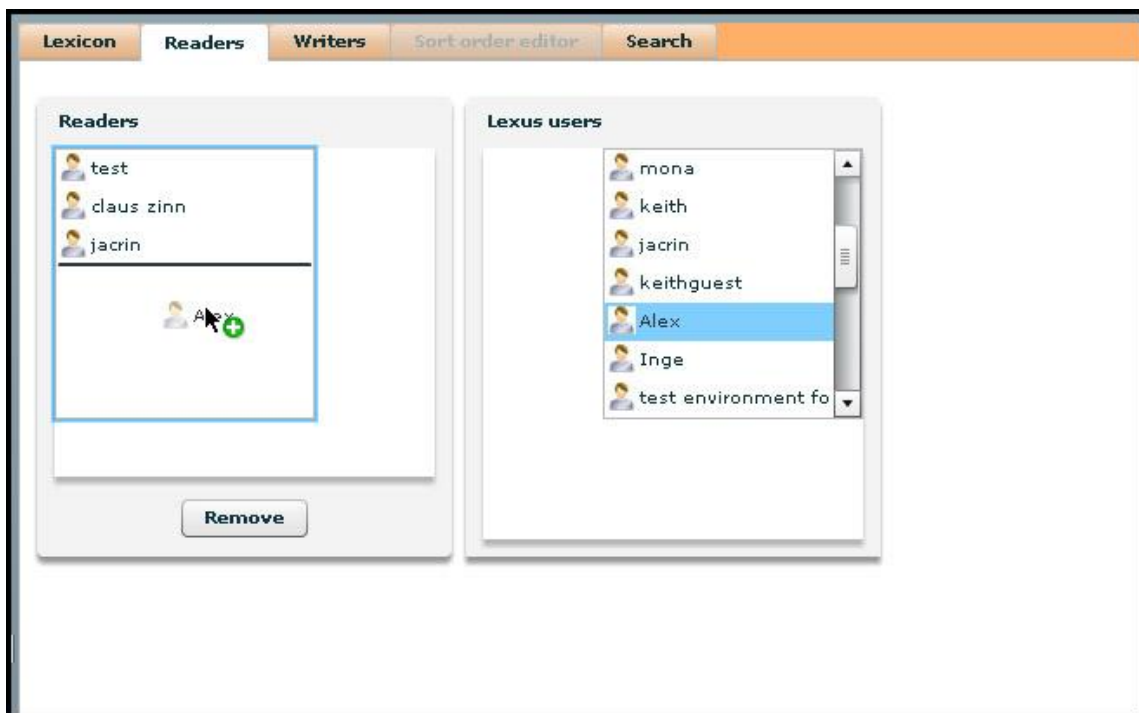


Figure 6.1. Assigning rights to other users

Chapter 7. ViCoS

ViCoS is an extension of LEXUS. It allows users to complement lexical spaces (as created by LEXUS) with ontological spaces. With ViCoS, users can define concepts they judge culturally relevant, and connect them to other concepts via a multitude of (user-definable) relation types. All concepts are anchored in the language to express them, and links can be established to objects in the multimedia archive to further describe them. The resulting conceptual space thus adds a new dimension to language documentation.

To switch from LEXUS to ViCoS use the menu option in the top frame of the Workspace Editor: LEXUS --> ViCoS.

ViCoS has its own manual. You can find it on the ViCoS web site [<http://www.lat-mpi.eu/tools/vicos/>].