Léacslann

Tutorial

Michal Boleslav Měchura
Fiontar, Dublin City University
27 July 2012

0 INTRODUCTION

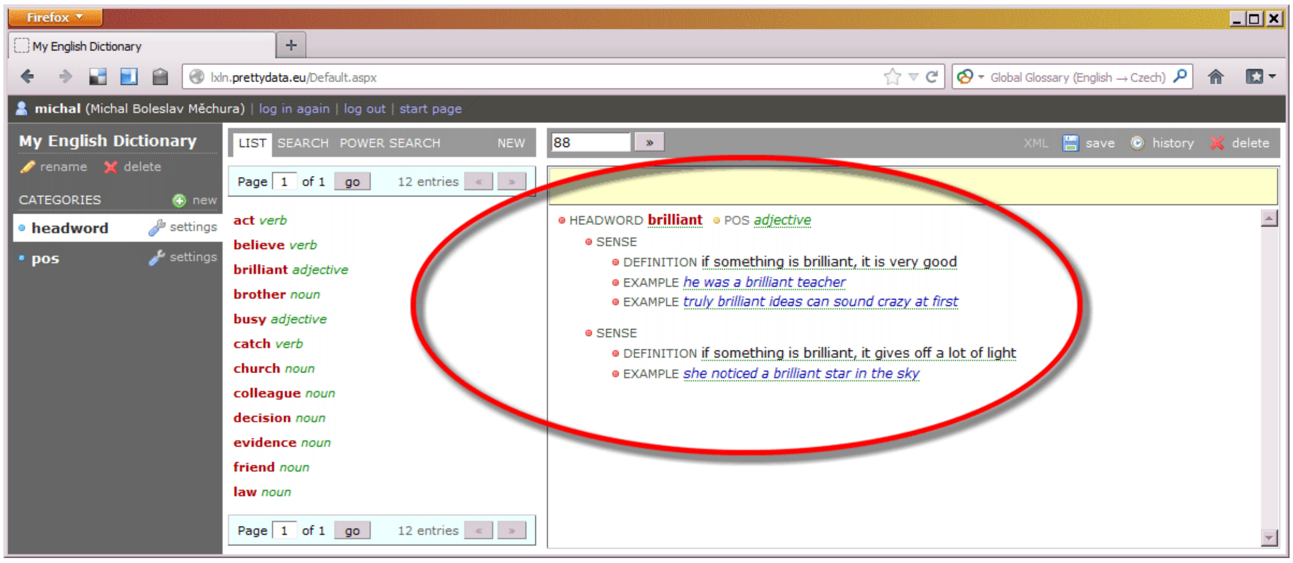
Léacslann is a web-based tool for editing **collections of structured hierarchical data**. You can use Léacslann to build a dictionary, a terminology database, an encyclopedia or indeed any collection of entries that have an arbitrary but strict internal structure. Léacslann is particularly suitable for applications in lexicography, terminography and reference science.

Léacslann is fully customizable. This means that you can tell Léacslann what structure you want your entries to have, and Léacslann will then give you the tools you need to compose entries that comply with that structure.

Léacslann is designed to be user-friendly: everything in Léacslann is done by clicking on things and typing into boxes. Even though you do need to understand a few concepts to use Léacslann successfully (these are all explained in this tutorial), you don't need to know or learn any programming languages.

Léacslann is multilingual. Even though Léacslann's user interface is in English, Léacslann can work with texts in any language.

0.1 What's the big idea?



Firefox
My English Dictionary
lxln.prettydata.eu/Default.aspx
Global Glossary (English → Czech)
michal (Michal Boleslav Měchura) | log in again | log out | start page
My English Dictionary
rename
delete
CATEGORIES
new
headword
settings
pos
settings
LIST
SEARCH
POWER SEARCH
NEW
Page 1 of 1 go 12 entries
88
»
XML
save
history
delete

act *verb*
believe *verb*
brilliant *adjective*
brother *noun*
busy *adjective*
catch *verb*
church *noun*
colleague *noun*
decision *noun*
evidence *noun*
friend *noun*
law *noun*

- HEADWORD **brilliant** • POS *adjective*
  - SENSE
    - DEFINITION if something is brilliant, it is very good
    - EXAMPLE *he was a brilliant teacher*
    - EXAMPLE *truly brilliant ideas can sound crazy at first*
  - SENSE
    - DEFINITION if something is brilliant, it gives off a lot of light
    - EXAMPLE *she noticed a brilliant star in the sky*

Figure 0.1: *An example entry*

Before we start using Léacslann, it is a good idea to spend some time reflecting about structured data in the abstract. We did say that Léacslann is a tool for editing collections of entries with an "arbitrary but strict" internal structure. Here's what we mean by that.

When you look at a dictionary, a terminology database or a similar collection of data, you will notice that it is made up of **entries** such as dictionary articles or terminological concepts, and that these entries all have a structure that looks like a **multi-levelled bulleted list** of elements. For example, a dictionary entry such as the one in Figure 0.1 consists of a headword – that's the top-level element – which branches out into one of more elements that represent senses of the headword. In turn, each sense can have a definition and one or more usage examples. Depending on the type of your dictionary, this structure can be elaborated further: a headword may need to have a part-of-speech label, a sense may be allowed to have various kinds of usage labels, a sense can have one or more translations for the headword, and so on.

It turns out that in the reference sciences (which encompass dictionary writing, terminology databases, encyclopedias and so on), all entries have such a structure. The details may differ, but each entry is always a multi-levelled bulleted list of elements. Depending on the type of collection you are working on, the types of elements allowed may be different (that's the "arbitrary" bit) but, once you've decided what those elements are, each entry must comply (that's the "strict" bit).

Léacslann allows you to specify what kinds of elements you want in your entries, and will then help you make sure that the entries you create will comply with that specification.

0.2 What's in this tutorial

Chapter 1 will start introducing Léacslann by showing you how to work with a collection of entries whose structure has already been specified. We will do this by looking at a simple monolingual dictionary like the one in Figure 1.

Chapter 2 will gently start taking you under the hood by showing you how the structure of the simple dictionary from Chapter 1 has been specified in Léacslann, and how you can alter and elaborate it in various ways. We will also show how to create your own stock from scratch ("stock" is Léacslann terminology for "a collection of entries").

Chapter 3 will explain how to use Léacslann's built-in search features, including its "power search" feature which allows you to interrogate your data in interesting and complex ways.

Chapter 4 will give you the back-story to Léacslann: how it stores its data, how it can be customized for enterprise-scale applications, who built Léacslann and why, how it's likely to be developed in the future, and under what circumstances you can use it for your own work.

0.3 How to access Léacslann

You can access Léacslann at this address: `http://lxln.prettydata.eu/` and log in with your user name and password. If you do not already have a user name, e-mail `valselob@gmail.com` and ask for one.

Once you've logged in, you will see three stocks that have been ready-made for you: *My English Dictionary*, *My Terminology Database* and *My Collection of Proverbs*. All three have been created to demonstrate various Léacslann features. The first one will be referred to often in this tutorial, while the remaining two are there for you to explore in your own time.

1 WORKING WITH DATA IN LÉACSLANN

In this chapter, we will demonstrate the editing features of Léacslann by taking you through a guided walking tour of one of the ready-made stocks, a simple monolingual dictionary titled *My English Dictionary*.

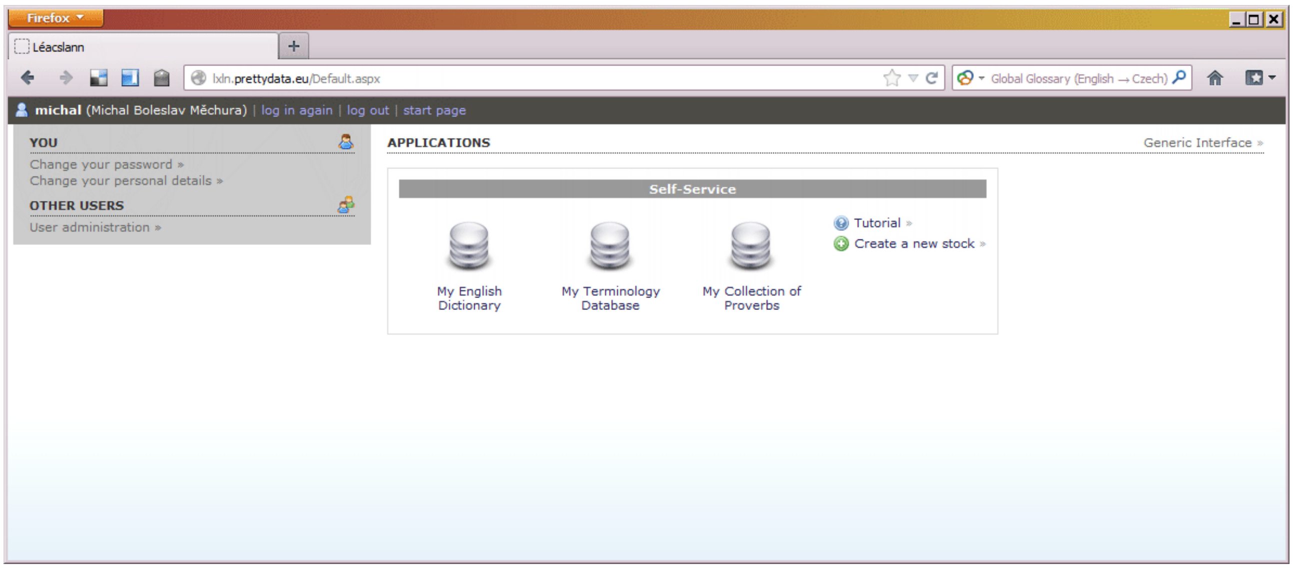
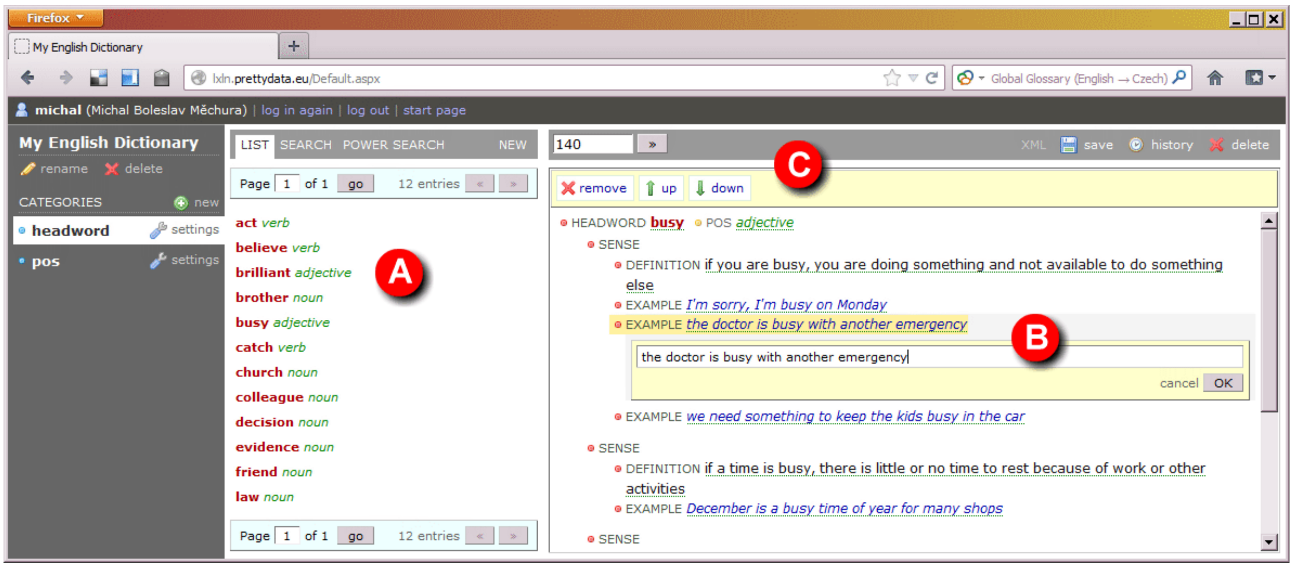


Figure 1.1: *Léacslann's start page*

1.1 Editing entries

After logging in to Léacslann, click on the *My English Dictionary* stock. You will see a screen similar to the one in Figure 1.2. This shows you the list of headwords currently included in the stock (A). There are only very few of them at the moment, but rest assured that Léacslann can comfortably handle hundreds of thousands of entries.



Firefox
My English Dictionary
lxln.prettydata.eu/Default.aspx
Global Glossary (English → Czech)
michal (Michal Boleslav Měchura) | log in again | log out | start page
My English Dictionary
rename
delete
CATEGORIES
new
headword
settings
pos
settings
LIST
SEARCH
POWER SEARCH
NEW
Page 1 of 1 go 12 entries
act verb
believe verb
brilliant adjective
brother noun
busy adjective
catch verb
church noun
colleague noun
decision noun
evidence noun
friend noun
law noun
Page 1 of 1 go 12 entries
A
140
XML
save
history
delete
C
remove
up
down
HEADWORD busy
POS adjective
SENSE
DEFINITION if you are busy, you are doing something and not available to do something else
EXAMPLE I'm sorry, I'm busy on Monday
EXAMPLE the doctor is busy with another emergency
the doctor is busy with another emergency
B
cancel
OK
EXAMPLE we need something to keep the kids busy in the car
SENSE
DEFINITION if a time is busy, there is little or no time to rest because of work or other activities
EXAMPLE December is a busy time of year for many shops
SENSE

Figure 1.2: *The 'My English Dictionary' stock, with one entry opened*

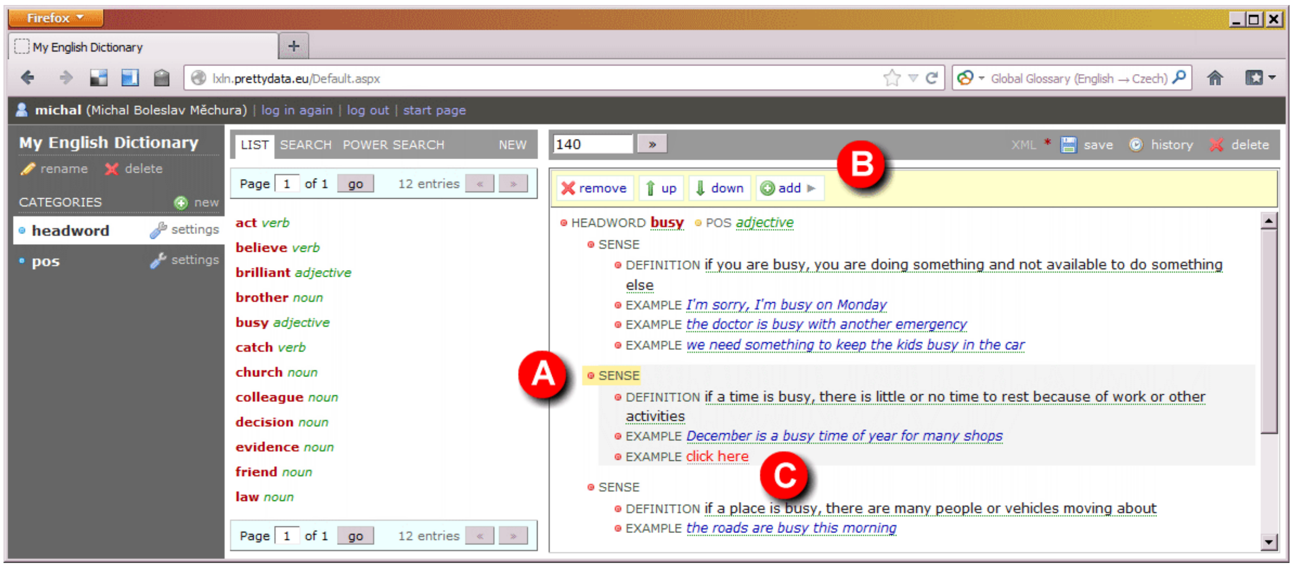
Click on a headword, such as *busy*, and you will see the complete dictionary article on the right. First of all, take some time to scrutinize what you see on screen. The article starts with a headword which has a part of speech (a "POS"). The headword is what's called an *element* and the part of speech is what's called an *attribute*. The *headword* element has several *sense* elements, and each *sense* element has one *definition* element and one or more *example* elements.

You can change any of the text by clicking on it (B). Click the OK button or press the Enter key on your keyboard to accept the changes. Click the Cancel link or click away from the text box to cancel editing.

Notice that each element and each attribute has two parts: a label, such as *sense* or *pos*, and a value (a piece of text you can click to edit). Some elements and attributes have both a label and a value, while some elements only have a label but no value (such as *sense*). If you click on a label (as opposed to a value), the element or attribute will become highlighted and some buttons will appear in the toolbar at the top (C). You can use these buttons to performs various operations on the element or attribute.

One thing you can do is move elements up and down. If you select one of the examples in the first sense of *busy*, you will be able to move it up and down the list. You can also delete the element altogether. You can do the same thing with senses: if you click on a sense label, you can move the sense up and down, while the entire contents of the sense moves with it. You can also delete a sense (and all its contents).

Notice that some elements can be deleted while others cannot. For example, Léacslann doesn't allow you to delete definitions, but does allow you to delete examples. This is because Léacslann knows that a sense is required to have a definition but that examples are optional. We will learn in Chapter in 2 how to make Léacslann "know" such things.



act *verb*
believe *verb*
brilliant *adjective*
brother *noun*
busy *adjective*
catch *verb*
church *noun*
colleague *noun*
decision *noun*
evidence *noun*
friend *noun*
law *noun*

HEADWORD **busy** POS *adjective*
- SENSE
  - DEFINITION if you are busy, you are doing something and not available to do something else
  - EXAMPLE *I'm sorry, I'm busy on Monday*
  - EXAMPLE *the doctor is busy with another emergency*
  - EXAMPLE *we need something to keep the kids busy in the car*
- SENSE
  - DEFINITION if a time is busy, there is little or no time to rest because of work or other activities
  - EXAMPLE *December is a busy time of year for many shops*
  - EXAMPLE click here
- SENSE
  - DEFINITION if a place is busy, there are many people or vehicles moving about
  - EXAMPLE *the roads are busy this morning*

Figure 1.3: *Adding a new element*

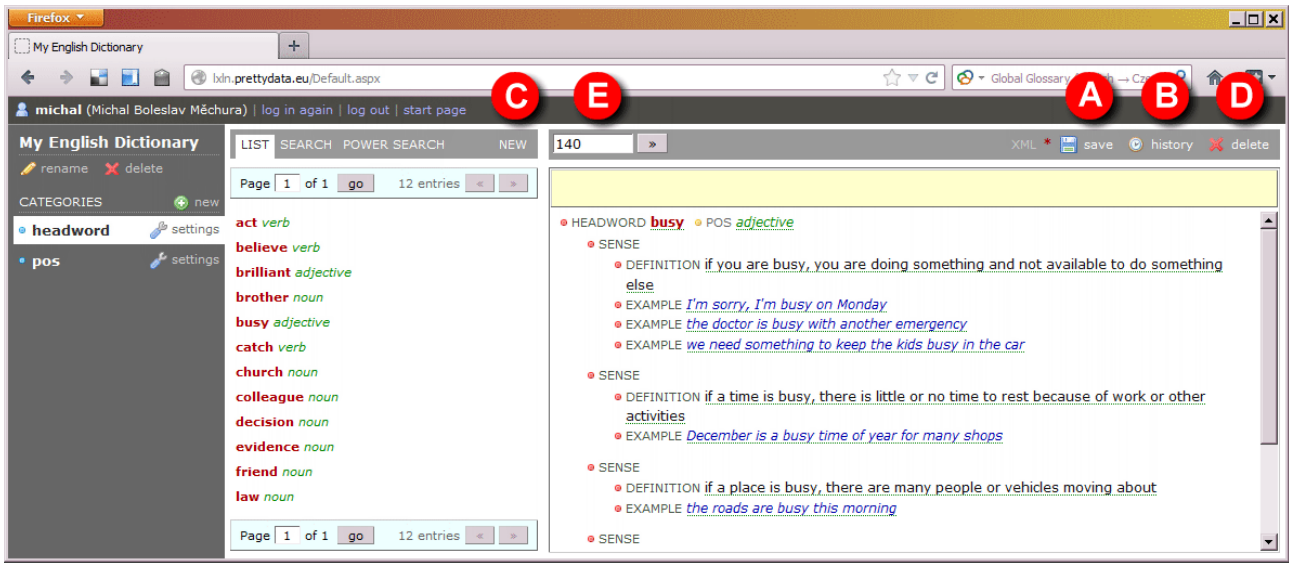
To add an example to a sense, select the *sense* label (Figure 1.3: A) and a button titled *add* will appear in the tool bar at the top (B). Click this button and an option to add an example will appear. Click this and a new example will be added to the sense (C). The example is empty, so you need to click the “click here” placeholder to add some text.

You will follow the same procedure to add a new sense to the headword. Select the headword, then click *add* and *sense*. Notice that when a new sense is added, it automatically contains an (empty) definition. This is because Léacslann knows that a sense must have a definition.

Notice also that Léacslann knows what kinds of “child” elements can be added to which “parent” element, so the contents of what you see after you click the *add* button is different based on what you have highlighted. Léacslann knows that a headword must have one or more senses, that a sense must have one definition, and that it may (but does not have to) have one or more examples. Léacslann also knows that some elements cannot have child elements at all, or that some elements cannot be deleted, and that’s why the *add* and *remove* buttons sometimes don’t appear at all.

By knowing all this, Léacslann only offers you the options that are relevant to the context in which you are. It also makes sure that the entry you are editing is always compliant with the structure that has been specified. We will learn in Chapter 2 how to specify structures.

1.2 Saving entries



act *verb*
believe *verb*
brilliant *adjective*
brother *noun*
busy *adjective*
catch *verb*
church *noun*
colleague *noun*
decision *noun*
evidence *noun*
friend *noun*
law *noun*

HEADWORD **busy** POS *adjective*
- SENSE
  - DEFINITION if you are busy, you are doing something and not available to do something else
  - EXAMPLE *I'm sorry, I'm busy on Monday*
  - EXAMPLE *the doctor is busy with another emergency*
  - EXAMPLE *we need something to keep the kids busy in the car*
- SENSE
  - DEFINITION if a time is busy, there is little or no time to rest because of work or other activities
  - EXAMPLE *December is a busy time of year for many shops*
- SENSE
  - DEFINITION if a place is busy, there are many people or vehicles moving about
  - EXAMPLE *the roads are busy this morning*
- SENSE

Figure 1.4: *Various elements in the user interface*

When you’ve made changes to an entry, you need to save it. To do that, click the *save* button at the top right (Figure 1.4: A). Any changes you have made by clicking and typing in the entry itself will not be saved until you click this button, so it’s very important not to forget to do that! To remind you, a small red star will appear next to the *save* button when you have made changes in the entry that have not been saved yet.

You will be pleased to hear that Léacslann keeps a record of every version of every entry you've ever saved. To see the complete changelog for an entry, click the *history* button (B). This will show you a chronological log of the entry's history, from the time it was created until the present (the most recent version is at the top). If you want to undo a change, you can re-instate a previous version by clicking the *restore* link next to it. This will return the entry to the state at which it was at that time.

1.3 Creating new entries

To create a new entry, click the *new* link at the top (C). A blank entry will be created, which you must now populate with data. Notice that some elements have been created automatically: because Léacslann knows that a headword must have a part-of-speech label and at least one sense, and that a sense must have a definition, it has created these things automatically for you.

Once you have populated the entry, you must save it by clicking the *save* button. Notice that the new entry will not appear automatically in the list of headwords on the left. You must click the *List* tab to reload the list.

1.4 Deleting entries

Obviously enough, you delete an entry by clicking the *delete* button at the top right (D). If you've deleted an entry by accident, you can resurrect it by going to its history and clicking the *recreate* link next to the most recent version.

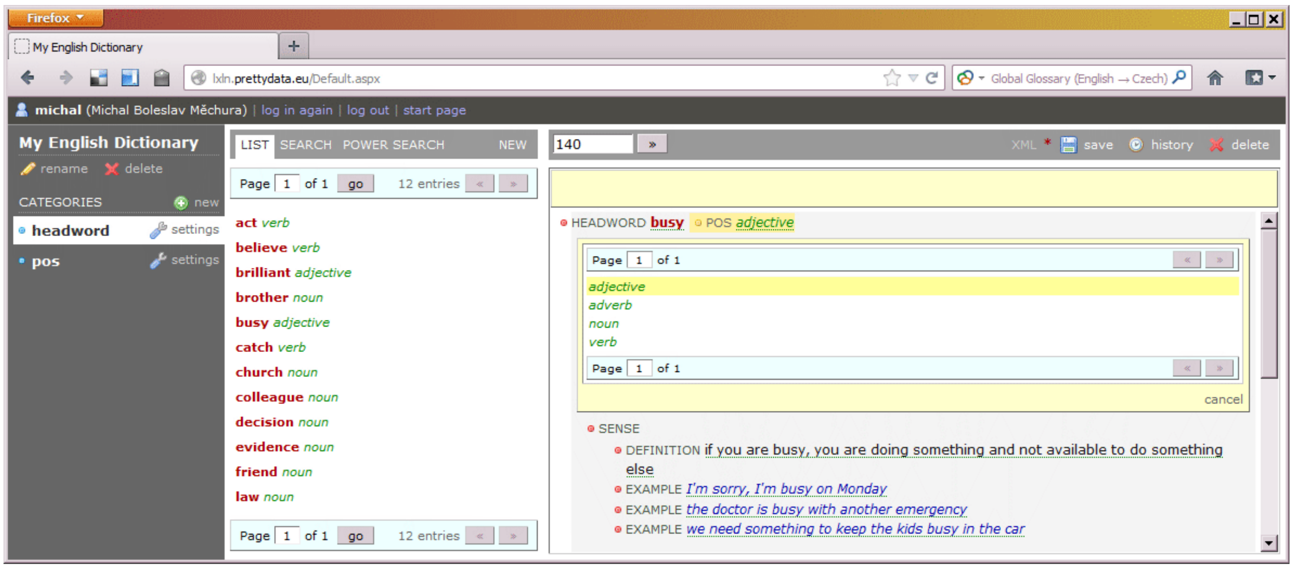
This is a good time to explain one important point about entries in Léacslann. Every entry in Léacslann has a unique number, which you can see in the little text box at top left (E). If you know the number of an entry, you can go straight to it by typing its number in the box.

Entry numbers in Léacslann are unique, which means that no two entries will ever have the same number. Even if an entry has been deleted, its number will never be used again for any other entry.

1.5 Working with references

Before we leave this chapter, we will return briefly to where we left off before we started looking at saving, creating and deleting entries. Open a headword such as *busy* and click on the value next to the part-of-speech label (Figure 1.5). You will notice that this time, you will get a list of values to choose from instead of a text box. The question to ask is, where do these options come from?

They come from a list which you will find if you click on the *pos* tab on the left-hand side (Figure 1.6). You will get a list of part-of-speech labels and the interesting thing is that you can edit, create a delete part-of-speech labels here in the same way as you did with headwords. In fact, as far as Léacslann cares, part-of-speech labels are entries just as headwords are entries – albeit entries with a much simpler structure, they only consist of one element with one value, while *headword* entries are much more “branchy”.



michal (Michal Boleslav Měchura) | log in again | log out | start page

My English Dictionary
rename delete
CATEGORIES new
headword settings
pos settings

LIST SEARCH POWER SEARCH NEW
Page 1 of 1 go 12 entries

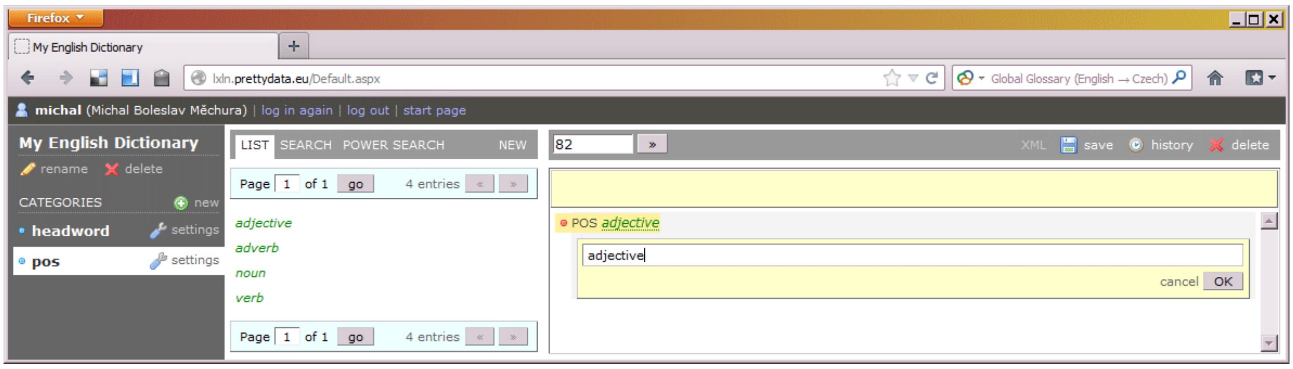
act *verb*
believe *verb*
brilliant *adjective*
brother *noun*
busy *adjective*
catch *verb*
church *noun*
colleague *noun*
decision *noun*
evidence *noun*
friend *noun*
law *noun*

Page 1 of 1 go 12 entries

140 » XML * save history delete

HEADWORD **busy** POS *adjective*
Page 1 of 1
adjective
adverb
noun
verb
Page 1 of 1
cancel
SENSE
DEFINITION if you are busy, you are doing something and not available to do something else
EXAMPLE *I'm sorry, I'm busy on Monday*
EXAMPLE *the doctor is busy with another emergency*
EXAMPLE *we need something to keep the kids busy in the car*

Figure 1.5: *A list of values to choose from*



michal (Michal Boleslav Měchura) | log in again | log out | start page

My English Dictionary
rename delete
CATEGORIES new
headword settings
pos settings

LIST SEARCH POWER SEARCH NEW
Page 1 of 1 go 4 entries

adjective
adverb
noun
verb

Page 1 of 1 go 4 entries

82 » XML save history delete

POS *adjective*
adjective
cancel OK

Figure 1.6: *Editing a part-of-speech label*

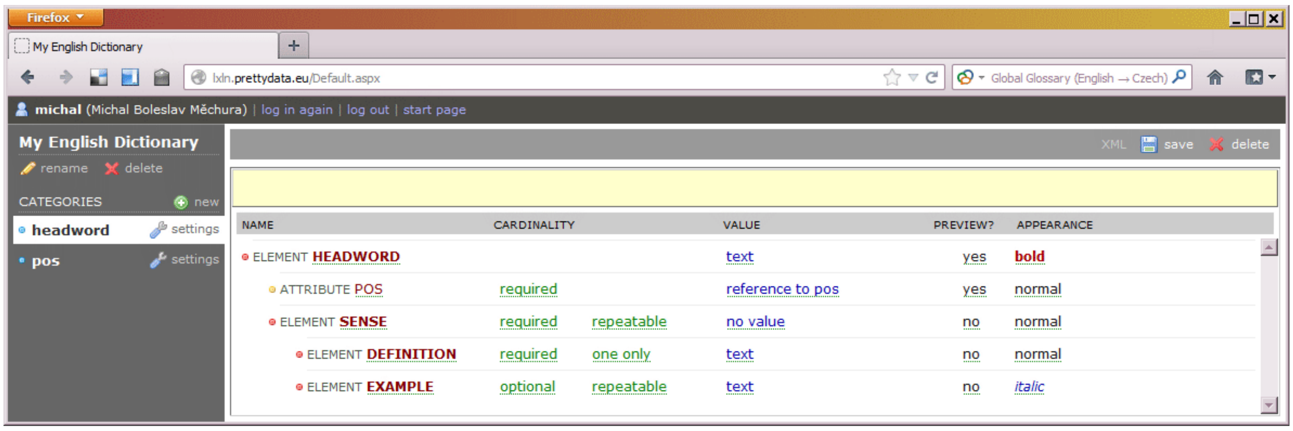
This is where Léacslann takes its list of part-of-speech labels from. Léacslann also knows that the value of the *pos* attribute of a *headword* element is to come from this list. In other words, what we have here is a (cross-)reference: a link from one entry (the headword) to another entry (the part-of-speech label).

What you see here is Léacslann’s ability to work with structures where an entry of some category contains a reference to another entry of a different category (or even to an entry of the same category). We will show you how to set up such structures in Chapter 2.

2 SPECIFYING DATA STRUCTURES

In this chapter, we will continue our guided walk though the *My English Dictionary* stock. We will show you how to understand and change the structure of entries. You now know that Léacslann somehow “knows” the structure that each entry is supposed to comply with. So far, we have been telling you that we will show you later where and how this structure is specified; this is what we will do now.

2.1 Understanding the structure of entries



Firefox
My English Dictionary
lxln.prettydata.eu/Default.aspx
Global Glossary (English → Czech)
michal (Michal Boleslav Měchura) | log in again | log out | start page
My English Dictionary
rename
delete
CATEGORIES
new
headword
settings
pos
settings
XML
save
delete

NAME	CARDINALITY		VALUE	PREVIEW?	APPEARANCE
ELEMENT HEADWORD			text	yes	bold
ATTRIBUTE POS	required		reference to pos	yes	normal
ELEMENT SENSE	required	repeatable	no value	no	normal
ELEMENT DEFINITION	required	one only	text	no	normal
ELEMENT EXAMPLE	optional	repeatable	text	no	*italic*

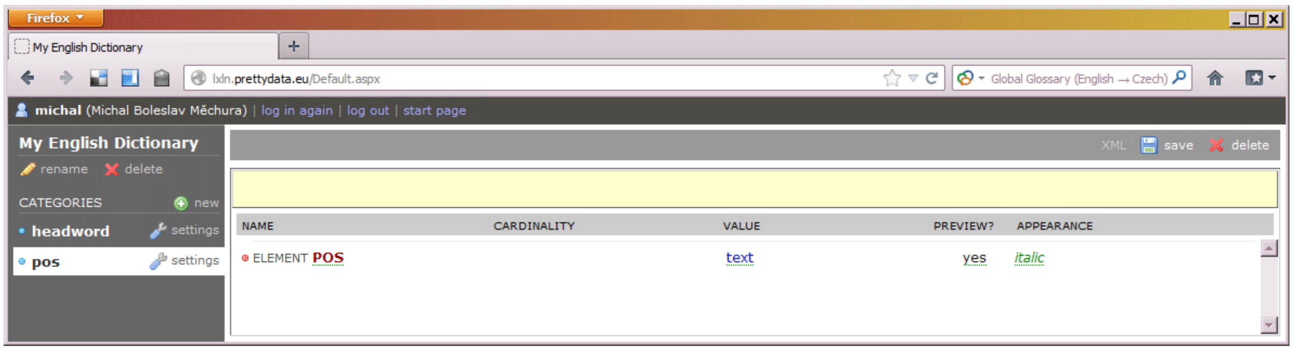
Figure 2.1: *The structure of ‘headword’ entries*

Click on the *settings* link inside the *headword* tab on the left-hand side of the screen. You will see a screen where the structure of *headword* entries is defined (Figure 2.1). Here is what it says:

- Each entry begins with an element called *headword* whose value is a string of text.
- Each *headword* element has an attribute called *pos* whose value is a reference to an entry of the *pos* category. Notice that the attribute is labelled as *required*, meaning that each headword must have one.
- Each *headword* element has one or more elements called *sense*. Notice that the sense element is labelled as *required* (meaning each sense must have one) and *repeatable* (meaning a sense may have more than one). Notice also that the *sense* element has no value.
- Inside each sense element, two further elements are allowed: *definition* (which is required and of which there must be only one) and *example* (which is optional and of which there may be more than one). They both have strings of text as their values.

This has explained what the columns *Name*, *Cardinality* and *Value* mean. You can ignore the remaining columns for now (*Preview* and *Appearance*), we will return to them later (in subchapter 2.5).

If you now click on the *settings* link next to the *pos* tab on the left-hand side, you will see that *pos* entries have much simpler structure (Figure 2.3): they're just one element with a text value. However, the idea is the same: each category has a structure specification and Léacslann makes sure that entries comply with it.

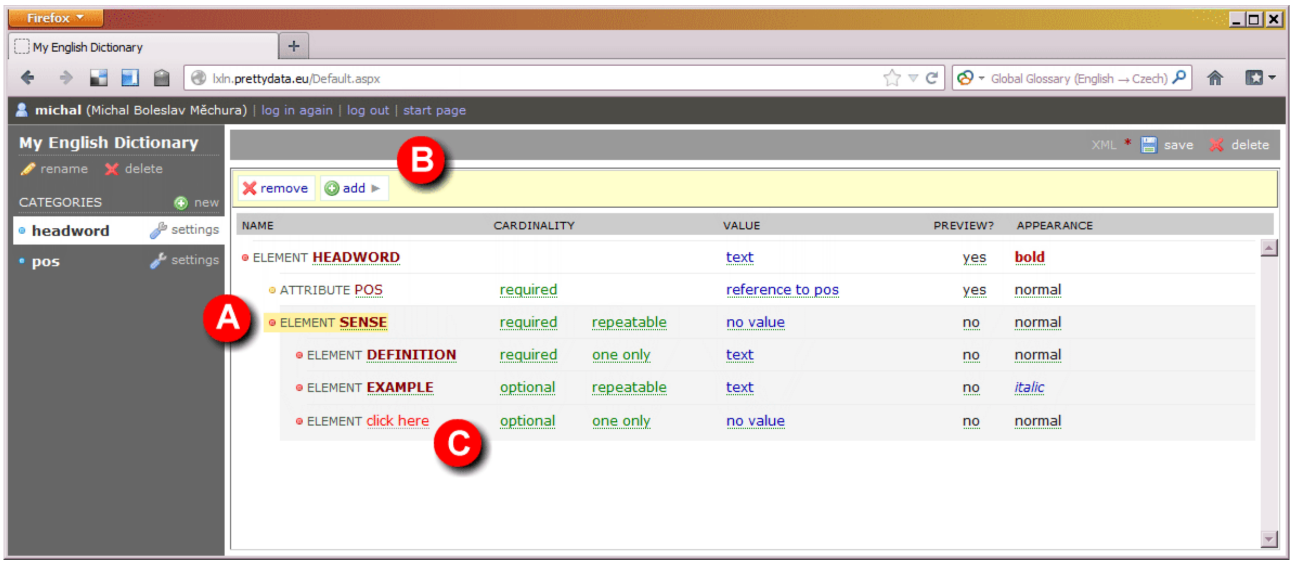

Firefox
My English Dictionary
lxln.prettydata.eu/Default.aspx
Global Glossary (English → Czech)
michal (Michal Boleslav Měchura) | log in again | log out | start page
My English Dictionary
rename
delete
CATEGORIES
new
headword
settings
pos
settings
XML
save
delete

NAME	CARDINALITY	VALUE	PREVIEW?	APPEARANCE
ELEMENT POS		text	yes	italic

Figure 2.2: *The structure of 'pos' entries*

2.2 Altering the structure of entries

We will now show you how to modify the structure of *headword* entries. Let's say you want to bilingualize your monolingual English dictionary; in other words, you want to add translations in another language, such as German or Arabic or Japanese. What you want is to be able to enter, for each sense, one or more translations of the headword. For that, you need to add a new child element to the *sense* element.


Firefox
My English Dictionary
lxln.prettydata.eu/Default.aspx
Global Glossary (English → Czech)
michal (Michal Boleslav Měchura) | log in again | log out | start page
My English Dictionary
rename
delete
CATEGORIES
new
headword
settings
pos
settings
XML *
save
delete
remove
add
A
B
C

NAME	CARDINALITY		VALUE	PREVIEW?	APPEARANCE
ELEMENT HEADWORD			text	yes	bold
ATTRIBUTE POS	required		reference to pos	yes	normal
ELEMENT SENSE	required	repeatable	no value	no	normal
ELEMENT DEFINITION	required	one only	text	no	normal
ELEMENT EXAMPLE	optional	repeatable	text	no	italic
ELEMENT click here	optional	one only	no value	no	normal

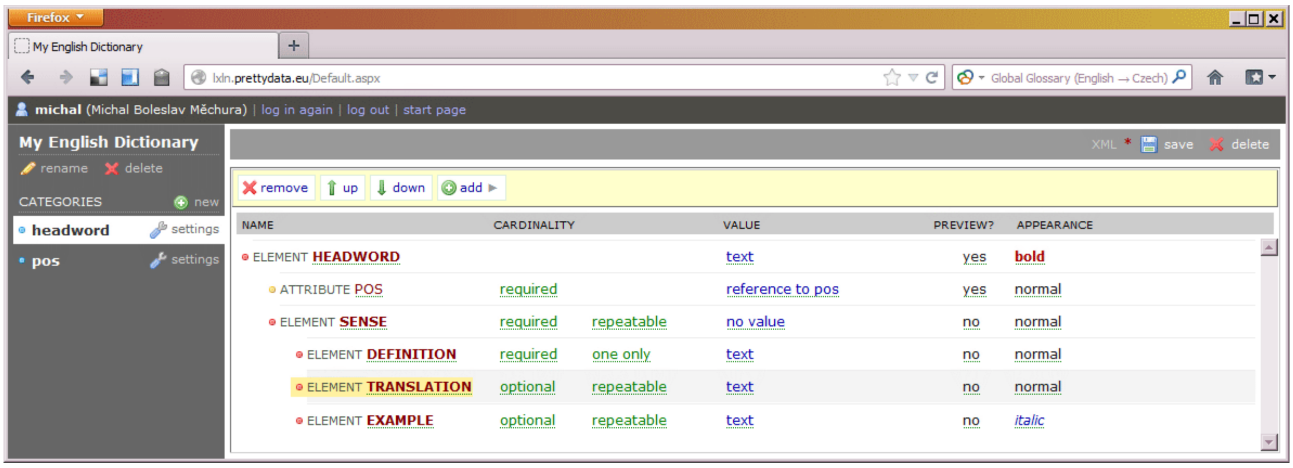
Figure 2.3: *Creating a new element*

Click on the *element* label of *sense* (Figure 2.3: A) and several buttons will appear in the tool bar at the top (B). Click *add*, then click *element*. This has created a new element (C) under *sense*. Click on the "click here" placeholder and give it a descriptive name, such as *translation*.

Let’s assume that you want translations to be optional (as opposed to required) and that you want to be able to include more than one under a sense. So, in the *Cardinality* column, leave the *optional* setting as it is and change the *one only* setting to *repeatable*.

Finally, you want each *translation* element to contain a string of text. Therefore, change the *Value* column from *no value* to *text*.

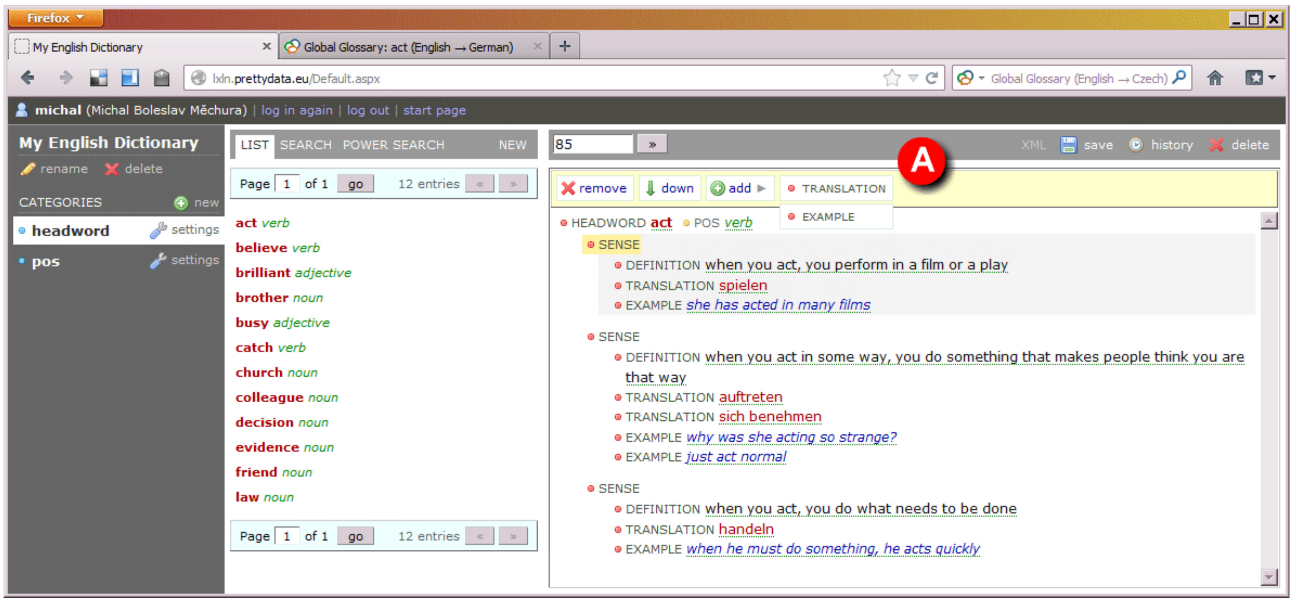
You probably want the translations to appear immediately after the definition and before any examples. So select the element you have just created (by clicking on its *element* label) and use the buttons in the tool bar to move it into a position between *definition* and *example*.


Firefox
My English Dictionary
lxln.prettydata.eu/Default.aspx
Global Glossary (English → Czech)
michal (Michal Boleslav Měchura) | log in again | log out | start page
My English Dictionary
rename
delete
CATEGORIES
new
headword
settings
pos
settings
XML
save
delete
remove
up
down
add

NAME	CARDINALITY		VALUE	PREVIEW?	APPEARANCE
ELEMENT HEADWORD			text	yes	bold
ATTRIBUTE POS	required		reference to pos	yes	normal
ELEMENT SENSE	required	repeatable	no value	no	normal
ELEMENT DEFINITION	required	one only	text	no	normal
ELEMENT TRANSLATION	optional	repeatable	text	no	normal
ELEMENT EXAMPLE	optional	repeatable	text	no	*italic*

Figure 2.4: *We have just created a new element*

The result should look as it does in Figure 2.4. Finally, save the changes you have just made by clicking the *save* button at the top right-hand side. Remember that changes have not been saved until you have clicked this button, so it is very important not to forget to do this. Just like in the previous chapter, a small red star next to the *save* button will remind you that have unsaved changes.



Page 1 of 1 go 12 entries

act *verb*
believe *verb*
brilliant *adjective*
brother *noun*
busy *adjective*
catch *verb*
church *noun*
colleague *noun*
decision *noun*
evidence *noun*
friend *noun*
law *noun*

Page 1 of 1 go 12 entries

HEADWORD **act** POS *verb*
- SENSE
  - DEFINITION when you act, you perform in a film or a play
  - TRANSLATION spielen
  - EXAMPLE *she has acted in many films*
- SENSE
  - DEFINITION when you act in some way, you do something that makes people think you are that way
  - TRANSLATION auftreten
  - TRANSLATION sich benehmen
  - EXAMPLE *why was she acting so strange?*
  - EXAMPLE *just act normal*
- SENSE
  - DEFINITION when you act, you do what needs to be done
  - TRANSLATION handeln
  - EXAMPLE *when he must do something, he acts quickly*

Figure 2.5: *Bilingualizing a dictionary*

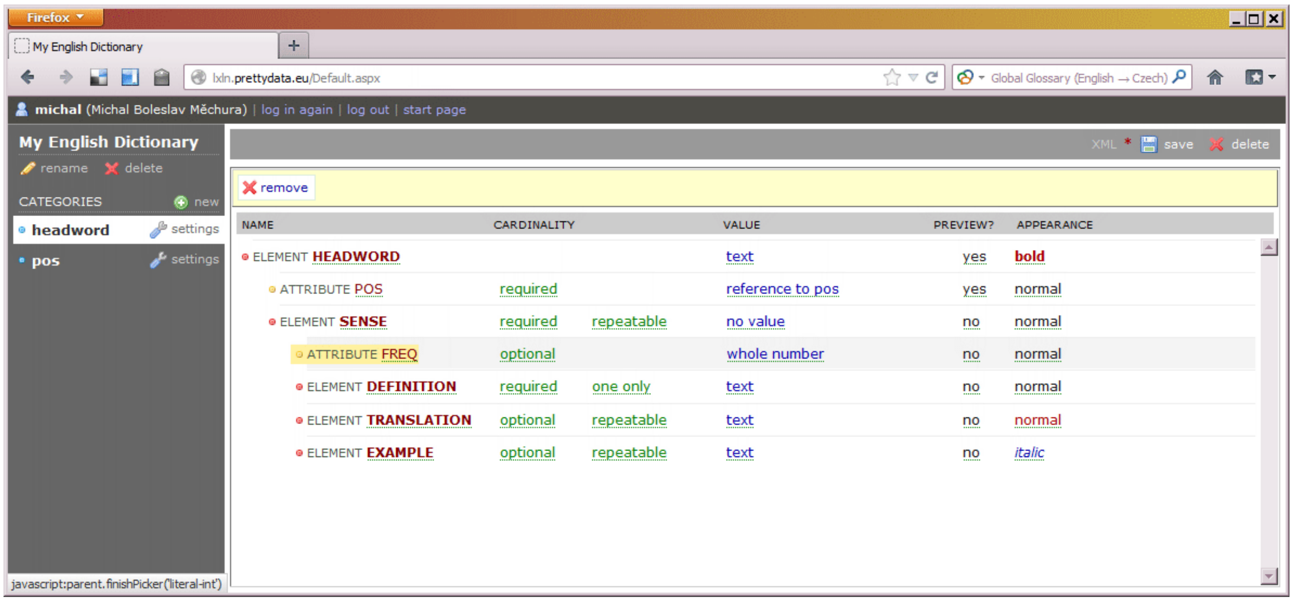
Once you have made and saved your changes, you can go back to entry editing by clicking the *headword* tab. If you now open a headword, highlight a *sense* element and click the *add* button, you will see that an additional option has appeared to add a *translation* element (Figure 2.5: A). You can now add translations to senses and bilingualize your dictionary!

If you want, you can add translations to *example* elements as well. Follow the same procedure as above, only make sure that you are adding the new element to *example* instead of *sense*. When deciding what to call this new element, you can call it *translation* again – Léacslann will not get confused if you have several elements with the same name. Just make sure that **you** don't get confused!

2.3 Working with different data types

When choosing the data type for your translation element, you may have noticed that there are more options than just *text*: there is *number*, *date*, *time* and many others. This shows you that Léacslann can handle many different data types – not just text – and we'll show you now what this can be useful for.

Let's assume that you want to add frequency data to your dictionary, in other words, data about how often each word is used. You could get data like this from a corpus, for example. Additionally, let's assume that you want your frequency data to be specified at the level of senses: you know how often each sense of each headword comes up in a corpus and you want to record these numbers in your dictionary.



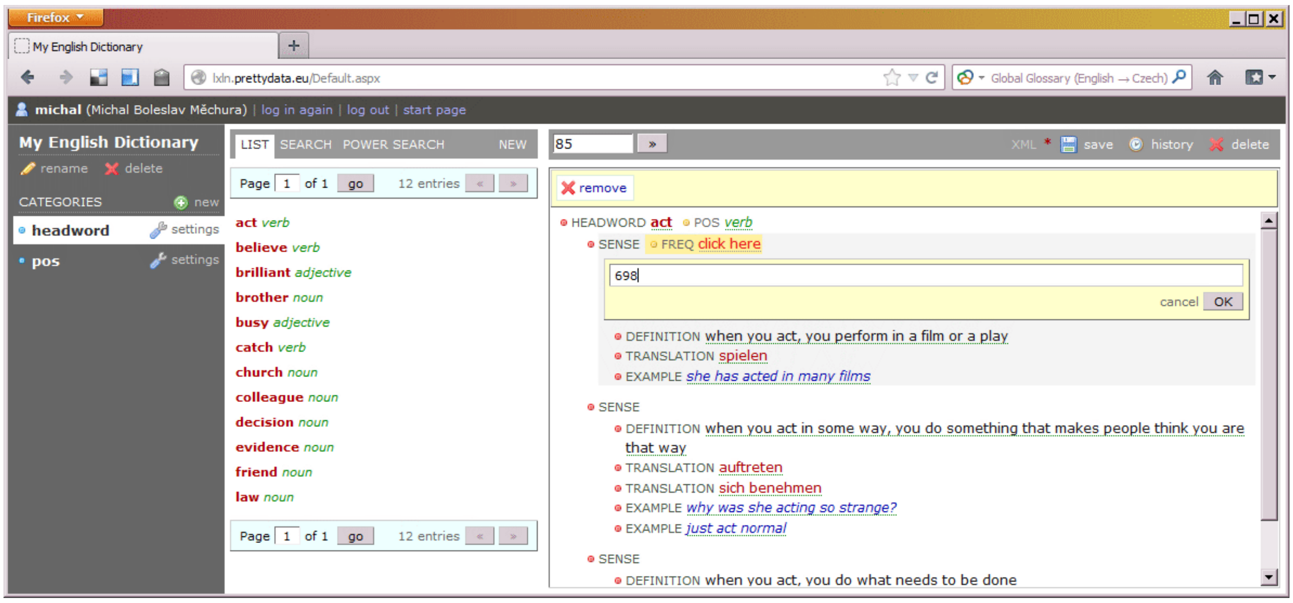
Firefox
My English Dictionary
lxln.prettydata.eu/Default.aspx
Global Glossary (English → Czech)
michal (Michal Boleslav Měchura) | log in again | log out | start page
My English Dictionary
rename
delete
CATEGORIES
new
headword
settings
pos
settings
XML
save
delete
remove
javascript:parent.finishPicker('literal-int')

NAME	CARDINALITY		VALUE	PREVIEW?	APPEARANCE
ELEMENT HEADWORD			text	yes	bold
ATTRIBUTE POS	required		reference to pos	yes	normal
ELEMENT SENSE	required	repeatable	no value	no	normal
ATTRIBUTE FREQ	optional		whole number	no	normal
ELEMENT DEFINITION	required	one only	text	no	normal
ELEMENT TRANSLATION	optional	repeatable	text	no	normal
ELEMENT EXAMPLE	optional	repeatable	text	no	italic

Figure 2.6: *Adding the 'freq' attribute*

The best way to go about this is to add an attribute called *freq* (or some such) to the *sense* element (Figure 2.6). Give this attribute the type *whole number* (a whole number is a number without a decimal point, for example 467 or 21,855). You may want to specify in the *Cardinality* column that this attribute will be *optional* (to give yourself the option of not supplying this data for senses whose frequency is unknown).

When you've done this and when you've saved your changes, you will now be able to add a *freq* attribute to each sense and fill it in with a number (Figure 2.7). If anybody attempts to fill it in with anything other than a number, Léacslann will display a warning message and will refuse to accept the value.



85

act *verb*
believe *verb*
brilliant *adjective*
brother *noun*
busy *adjective*
catch *verb*
church *noun*
colleague *noun*
decision *noun*
evidence *noun*
friend *noun*
law *noun*

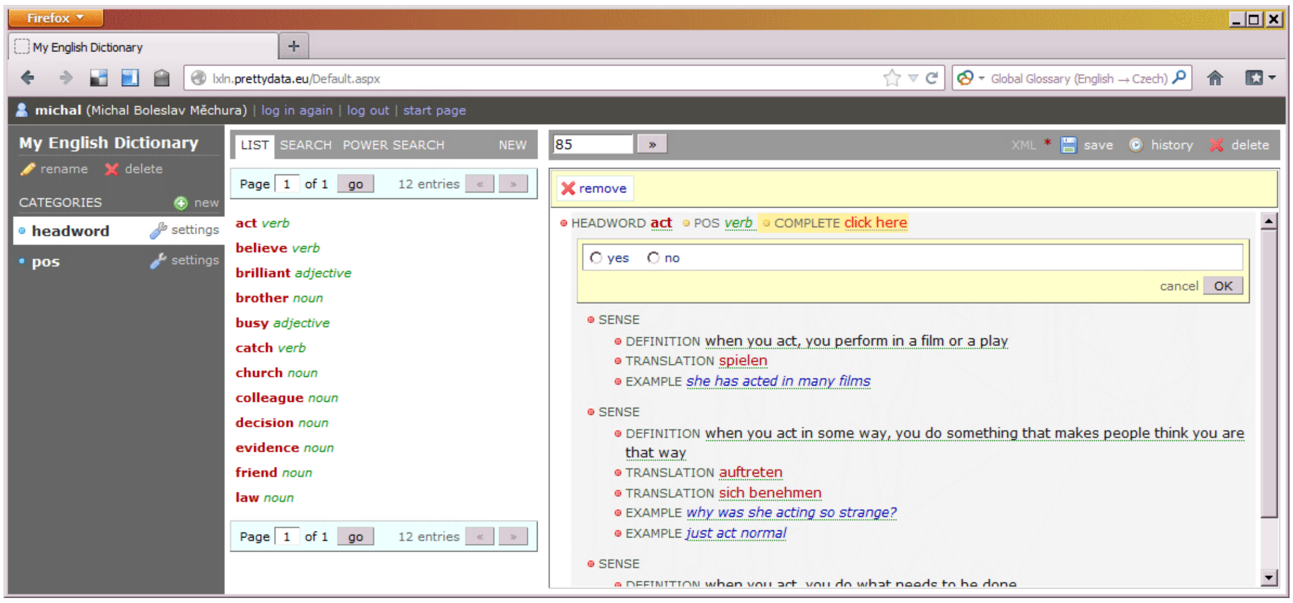
HEADWORD act POS *verb*
SENSE FREQ click here
698
DEFINITION when you act, you perform in a film or a play
TRANSLATION spielen
EXAMPLE *she has acted in many films*
SENSE
DEFINITION when you act in some way, you do something that makes people think you are that way
TRANSLATION auftreten
TRANSLATION sich benehmen
EXAMPLE *why was she acting so strange?*
EXAMPLE *just act normal*
SENSE
DEFINITION when you act, you do what needs to be done

Figure 2.7: *Entering data into the 'freq' attribute*

This has shown you how to create attributes and elements that contain numbers. For the sake of more exercise, let's take a look at another data type, a data type called *yes/no*. You can use this to create elements and attributes whose values can have one of only two values, *yes* or *no*. Let's assume you want to flag whether each headword is complete or whether it is work in progress. You can do this by adding to the *headword* element an attribute called *complete* or some such, and giving it the data type *yes/no*. You will then be able to record for each headword whether it is complete or not (Figure 2.8). We will leave it up to you create this attribute in your own time if you want more practice.

In theory, you do not have to bother with data types and always only use *text*. There is nothing to stop you from assigning the data type *text* to the *freq* and *complete* attributes, and to record all data as text. For example, you could express the completion status of a headword by typing the words "yes" or "no". However, this is not recommended. There are advantages to telling Léacslann what type of data is supposed to go into each element and attribute. One advantage is that Léacslann will make sure that the data you fill in always complies with the data type you had specified: it will only accept *yes* or *no* and nothing else. This ensures consistency. After all, remember that Léacslann is a tool for working with "arbitrary but **strict**" structures!

When we start looking at Léacslann's search features in Chapter 3, we will show you how you can use data types such as *whole number* and *yes/no* to interrogate your data in interesting ways. For example, you can search for headwords that are labelled as incomplete, headwords that have a sense with a frequency higher than a given number, and so on.



Firefox
My English Dictionary
lxln.prettydata.eu/Default.aspx
Global Glossary (English → Czech)
michal (Michal Boleslav Měchura) | log in again | log out | start page

My English Dictionary
rename delete
CATEGORIES new
headword settings
pos settings

LIST SEARCH POWER SEARCH NEW
Page 1 of 1 go 12 entries

act *verb*
believe *verb*
brilliant *adjective*
brother *noun*
busy *adjective*
catch *verb*
church *noun*
colleague *noun*
decision *noun*
evidence *noun*
friend *noun*
law *noun*

Page 1 of 1 go 12 entries

85 » XML * save history delete

remove
HEADWORD act POS verb COMPLETE click here
yes no
cancel OK
SENSE
DEFINITION when you act, you perform in a film or a play
TRANSLATION spielen
EXAMPLE *she has acted in many films*
SENSE
DEFINITION when you act in some way, you do something that makes people think you are that way
TRANSLATION auftreten
TRANSLATION sich benehmen
EXAMPLE *why was she acting so strange?*
EXAMPLE *just act normal*
SENSE
DEFINITION when you act, you do what needs to be done

Figure 2.8: *Entering a value into the 'complete' attribute*

2.4 Creating structures with references

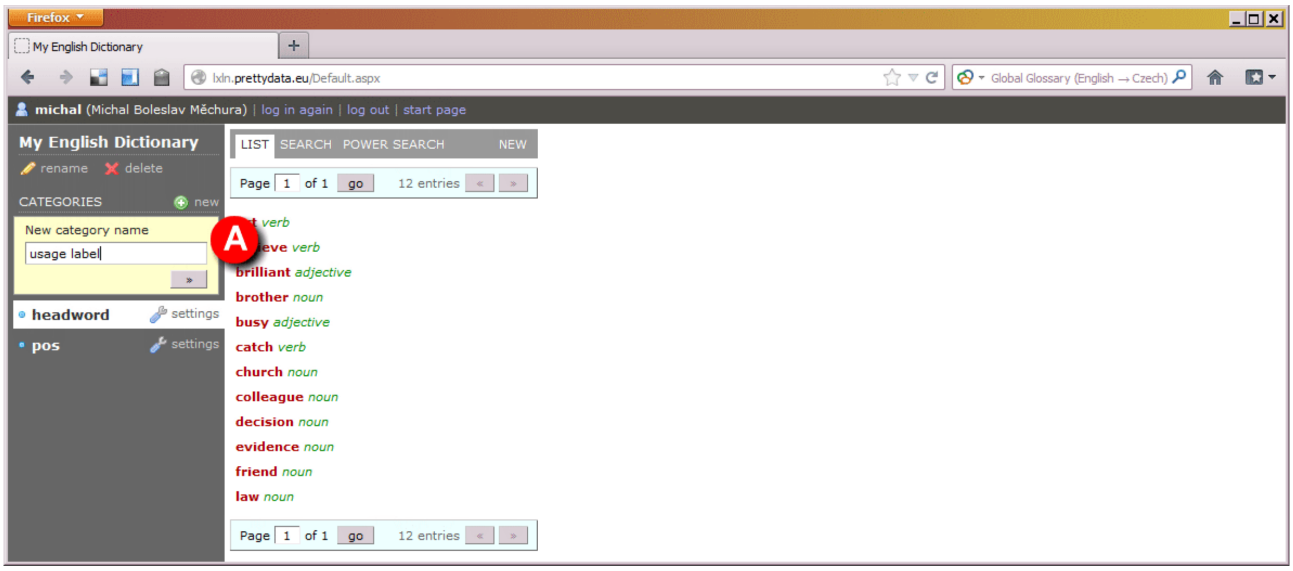
One thing we didn't cover in our discussion of data types in the previous subchapter is reference data types; that is, elements and attributes whose values are references to other entries in the stock. You already know that the *pos* attribute of the *headword* element is such an attribute: it contains a reference to entry of the *pos* category.

We will now show how to set such a thing up on your own. As usual, we will do this with an example. Let's assume you want to be able to include usage labels in your dictionary: labels such as *formal*, *informal*, *archaic*, *neologism*, *vulgar* and so on. You want to be able to assign these labels to senses.

The first thing you need to do is create a list of usage labels – and, to be able to do that, you need to create a category for them. So, click the *new* link next to the *Categories* heading on the left-hand side of your screen and type a descriptive name for the category, for example *usage label* (Figure 2.9: A).

The internal structure of usage labels will be very simple and will be more or less identical to that of part-of-speech labels: it'll just be a single string of text. So, change the data type of the *usage label* element from *no value* to text (Figure 2.10). Don't forget to save this change by clicking the *save* button.

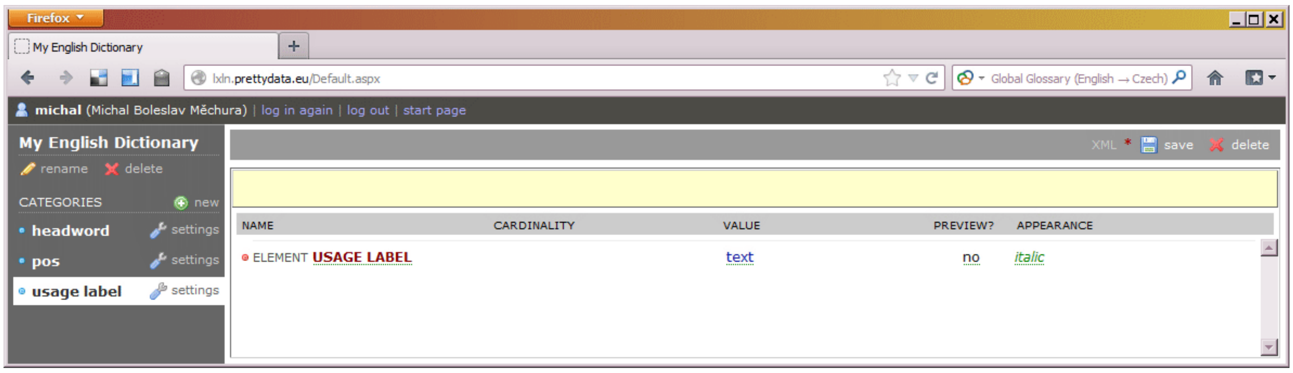
Now that the *usage label* category has been created, you need to create a few *usage label* entries. This is easy to do and you already know how to do this. Click on the usage label tab and add a few labels such as *formal*, *informal* and whatever else you think you might need in your dictionary (Figure 2.11). (For more guidance on adding entries, look back at the subchapter *1.3 Editing entries.*)


Firefox
My English Dictionary
lxln.prettydata.eu/Default.aspx
Global Glossary (English → Czech)
michal (Michal Boleslav Měchura) | log in again | log out | start page
My English Dictionary
rename
delete
CATEGORIES
new
LIST
SEARCH
POWER SEARCH
NEW
A
headword
settings
pos
settings
Page 1 of 1 go 12 entries

New category name
usage label

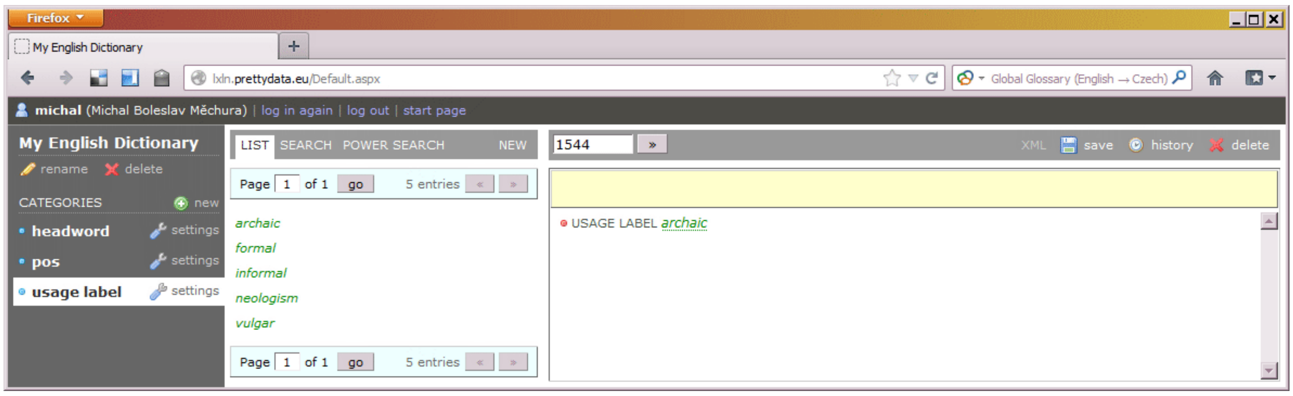
t verb
eve verb
brilliant adjective
brother noun
busy adjective
catch verb
church noun
colleague noun
decision noun
evidence noun
friend noun
law noun

Figure 2.9: *Creating a new category*


Firefox
My English Dictionary
lxln.prettydata.eu/Default.aspx
Global Glossary (English → Czech)
michal (Michal Boleslav Měchura) | log in again | log out | start page
My English Dictionary
rename
delete
CATEGORIES
new
headword
settings
pos
settings
usage label
settings
XML
save
delete

NAME	CARDINALITY	VALUE	PREVIEW?	APPEARANCE
ELEMENT USAGE LABEL		text	no	italic

Figure 2.10: *The structure of the 'usage label' category*

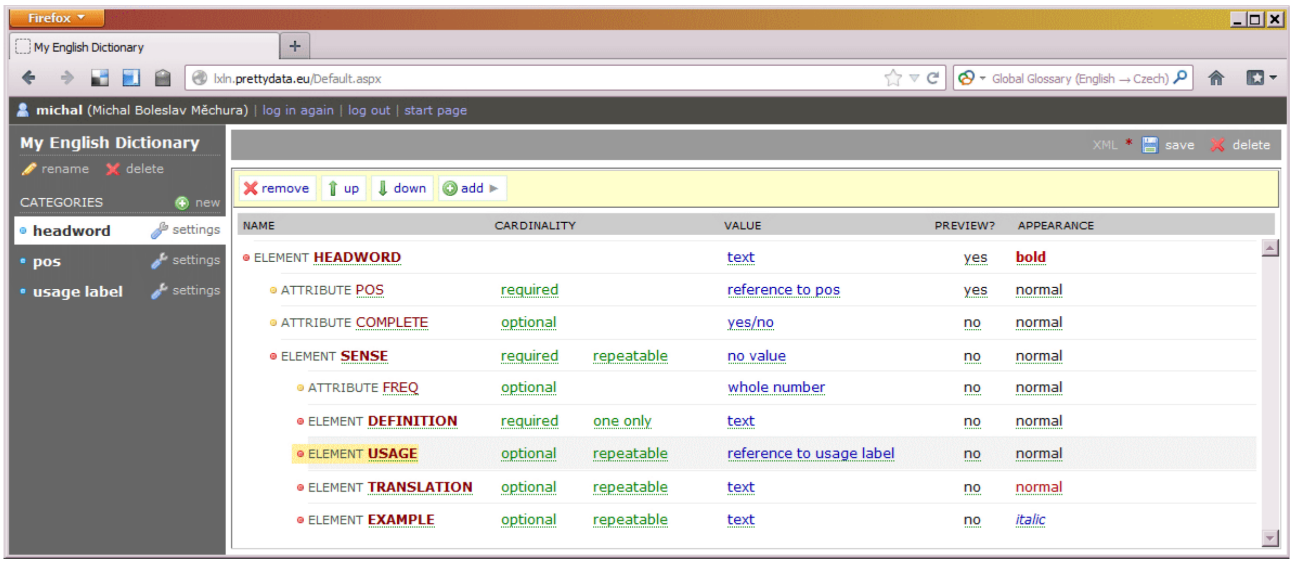


archaic
formal
informal
neologism
vulgar

USAGE LABEL archaic

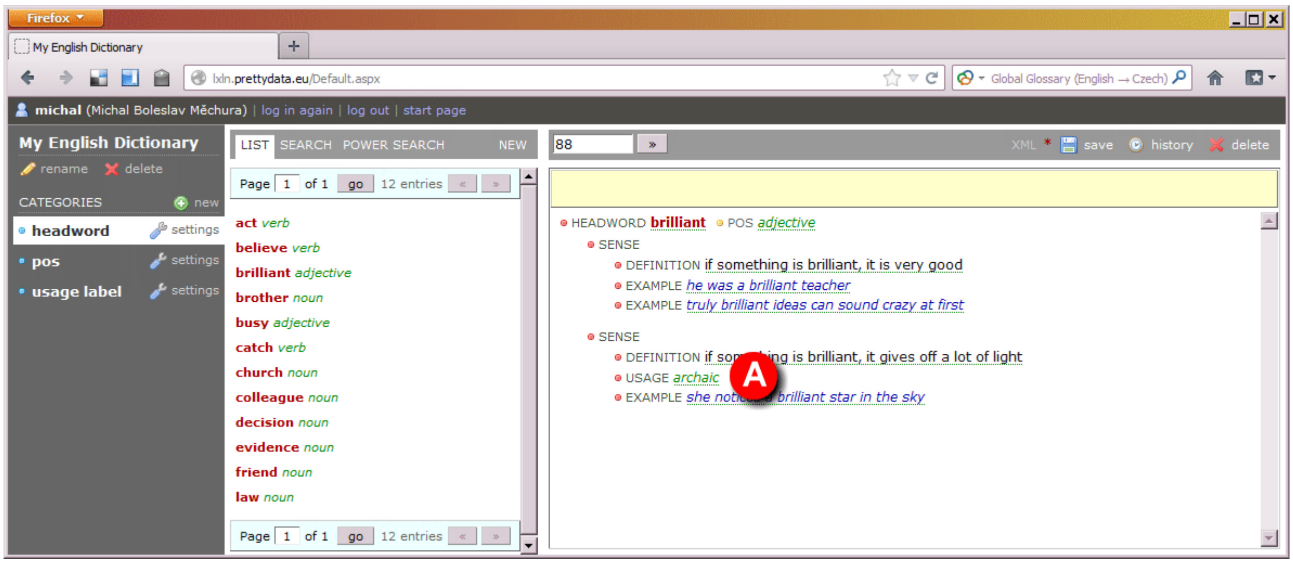
Figure 2.11: *We have just created a few 'usage label' entries*

Now that you have a list of usage labels in your stock, the next thing you need to do is update the structure of *headword* entries. Click the *settings* link in the *headword* tab and highlight the *sense* element. Add a new child element to *sense*, give it a name (such as *usage*) and change its data type from *no value* to *reference to usage label*. You probably want usage labels to be *optional* and *repeatable* (so that you can enter none, one or more than one), so change the *Cardinality* column accordingly. Last but not least, you probably want these labels to appear the beginning of a sense rather than at the end, so move the element up to the beginning, right after *definition* (or even before it if you prefer). The result should look as in Figure 2.12. Finally, don't forget to save your changes.



NAME	CARDINALITY		VALUE	PREVIEW?	APPEARANCE
ELEMENT HEADWORD			text	yes	bold
ATTRIBUTE POS	required		reference to pos	yes	normal
ATTRIBUTE COMPLETE	optional		yes/no	no	normal
ELEMENT SENSE	required	repeatable	no value	no	normal
ATTRIBUTE FREQ	optional		whole number	no	normal
ELEMENT DEFINITION	required	one only	text	no	normal
ELEMENT USAGE	optional	repeatable	reference to usage label	no	normal
ELEMENT TRANSLATION	optional	repeatable	text	no	normal
ELEMENT EXAMPLE	optional	repeatable	text	no	italic

Figure 2.12: *Creating a reference to 'usage label' entries from 'headword' entries*



Page 1 of 1 go 12 entries

act *verb*
believe *verb*
brilliant *adjective*
brother *noun*
busy *adjective*
catch *verb*
church *noun*
colleague *noun*
decision *noun*
evidence *noun*
friend *noun*
law *noun*

Page 1 of 1 go 12 entries

HEADWORD **brilliant** POS *adjective*
SENSE
DEFINITION if something is brilliant, it is very good
EXAMPLE *he was a brilliant teacher*
EXAMPLE *truly brilliant ideas can sound crazy at first*
SENSE
DEFINITION if so[illegible]ing is brilliant, it gives off a lot of light
USAGE *archaic*
EXAMPLE *she noti[illegible] brilliant star in the sky*

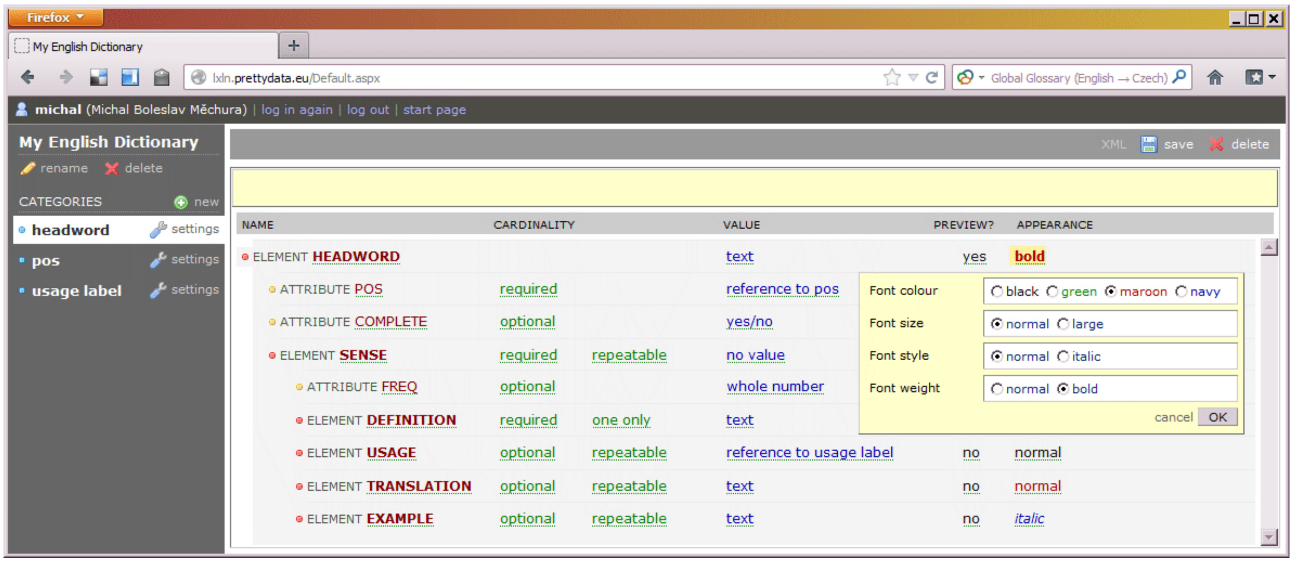
Figure 2.13: *A usage label has been added to a 'headword' entry*

You have now achieved what you wanted: you can now add usage labels to senses. Try opening a *headword* entry and try adding one or more usage labels to a sense (Figure 2.13: A). To review, what we have done is this: first we have created a new category for usage labels, then we have populated the category with a few entries, finally we have added an element to the structure of *headword* entries whose data type is *reference to usage label*. This is the process you will follow each time you want to have an element in your data whose value comes from a closed list of values: a list of usage labels, part-of-speech labels, language names or whatever else.

You can follow the same process to set up cross-references between entries of the same type. Let's say you want to be able include a cross-reference from a *sense* of a *headword* to another *headword*. This is easy to do, you simply need to add an element to the end of *sense* called *see also* or some such and set its data type to *reference to headword*. We will leave that for you to try out in your own time.

There is one thing we have left unexplained in our discussion. When we were updating the structure of *headword* entries to include a reference to *usage label*, we decided to create an element rather than an attribute. This begs the question, what is the difference between an element and attribute? The only difference is that elements can be *repeatable* while attributes cannot. In other words, a parent element can have more than one child element of the same name (if you set it as *repeatable*) but it can only have one attribute of the same name. The other difference is that attributes are displayed on the same line of text as the element they belong to, while child element are displayed on a new (indented) line underneath, which makes for a more compact layout. So, when choosing between element and attribute, you need to take these things into consideration. *(We are aware that the distinction between elements and attributes is a bit arcane and unhelpful, and we are considering getting rid of attributes altogether in a future version of Léacslann.)*

2.5 Formatting entries


Firefox
My English Dictionary
lxln.prettydata.eu/Default.aspx
Global Glossary (English → Czech)
michal (Michal Boleslav Měchura) | log in again | log out | start page
My English Dictionary
rename
delete
CATEGORIES
new
headword
settings
pos
settings
usage label
settings
XML
save
delete

NAME	CARDINALITY		VALUE	PREVIEW?	APPEARANCE
ELEMENT **HEADWORD**			text	yes	**bold**
ATTRIBUTE POS	required		reference to pos		
ATTRIBUTE COMPLETE	optional		yes/no		
ELEMENT **SENSE**	required	repeatable	no value		
ATTRIBUTE FREQ	optional		whole number		
ELEMENT **DEFINITION**	required	one only	text		
ELEMENT **USAGE**	optional	repeatable	reference to usage label	no	normal
ELEMENT **TRANSLATION**	optional	repeatable	text	no	normal
ELEMENT **EXAMPLE**	optional	repeatable	text	no	*italic*

Font colour	black green [✓ maroon] navy
Font size	[✓ normal] large
Font style	[✓ normal] italic
Font weight	normal [✓ bold]

cancel OK

Figure 2.14: *The 'Preview' and 'Appearance' columns*

You can congratulate yourself because you have now covered all the difficult bits. You now know how to create structures in Léacslann and how to populate them with data. All that remains to look at is formatting settings which you will find in the last two columns (titled *Preview* and *Appearance*) on the *settings* screen (Figure 2.14).

The *Appearance* column determines what the value will look like on screen. You probably want things like headwords and translation equivalents to stand out on screen, while other data should be less prominent. So it's a good idea to play around with the formatting options to emphasize certain content. On the other hand, it's an equally good idea not to overdo it because, when everything is emphasized, nothing is.

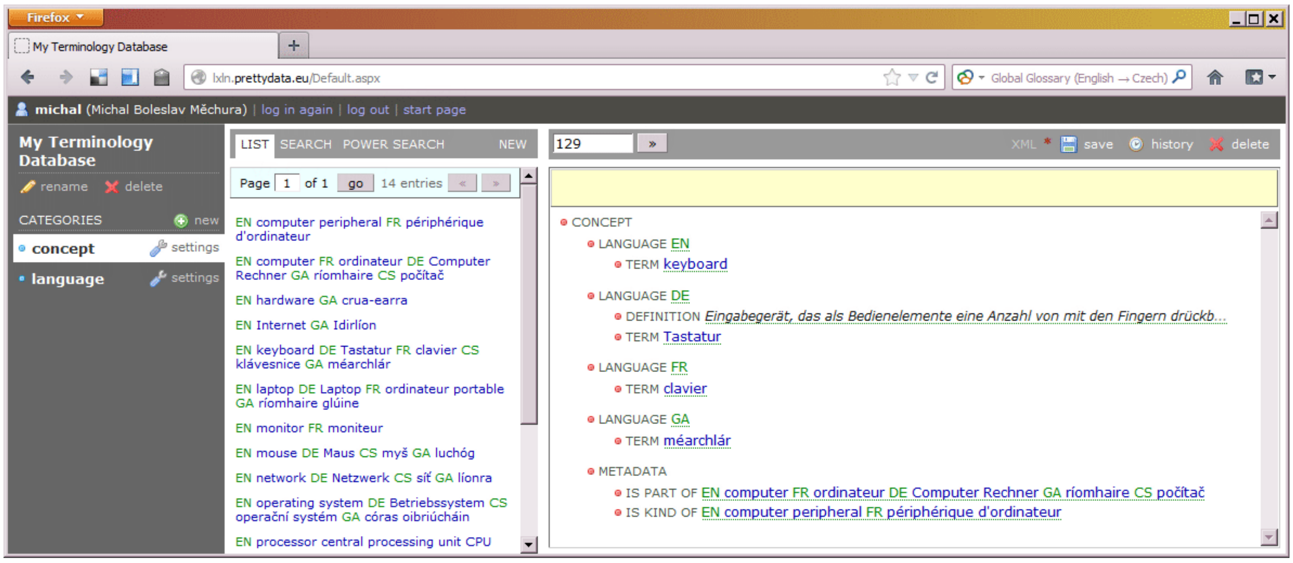
The *Preview* column determines whether the element or attribute will appear as part of the entry preview that you see in the listing pane in the middle of the screen. You do not have to worry about this very much because, if you forget to set it to *yes* for any element or attribute at all, Léacslann will try to find the most suitable ones automatically and will set them to *yes* for you.

When entry previews are listed in the listing pane in the middle of the screen, they are sorted alphabetically based in the contents of the preview. *(At the moment, Léacslann sorts using the same collation for all entries, regardless of language. This means that the alphabetical sorting order may be slightly incorrect for some languages. We are planning to introduce more sophistication into this in a future version of Léacslann.)*

2.6 A brief look at other stocks

So far in this tutorial we have been looking at the *My English Dictionary* stock, which has served us well to demonstrate all the important features of Léacslann. But Léacslann can be used for much more than just dictionaries, and we will demonstrate that now by giving a brief overview of the other two stocks that are ready-made for you in Léacslann: *My Terminology Database* and *My Collection of Proverbs*.

2.6.1 My Terminology Database


Firefox
My Terminology Database
lxln.prettydata.eu/Default.aspx
Global Glossary (English → Czech)
michal (Michal Boleslav Měchura) | log in again | log out | start page
rename
delete
CATEGORIES
new
settings
LIST
SEARCH
POWER SEARCH
NEW
go
XML
save
history
delete

My Terminology Database

concept

language

Page 1 of 1 — 14 entries

EN computer peripheral FR périphérique d'ordinateur
EN computer FR ordinateur DE Computer Rechner GA ríomhaire CS počítač
EN hardware GA crua-earra
EN Internet GA Idirlíon
EN keyboard DE Tastatur FR clavier CS klávesnice GA méarchlár
EN laptop DE Laptop FR ordinateur portable GA ríomhaire glúine
EN monitor FR moniteur
EN mouse DE Maus CS myš GA luchóg
EN network DE Netzwerk CS síť GA líonra
EN operating system DE Betriebssystem CS operační systém GA córas oibriúcháin
EN processor central processing unit CPU

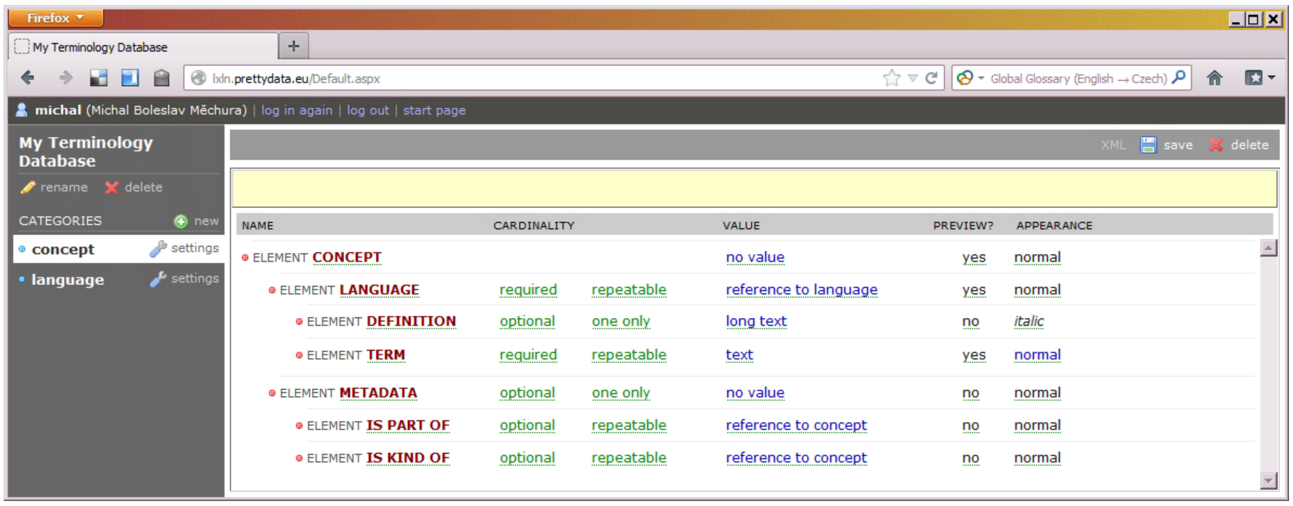
129

- CONCEPT
  - LANGUAGE EN
    - TERM keyboard
  - LANGUAGE DE
    - DEFINITION *Eingabegerät, das als Bedienelemente eine Anzahl von mit den Fingern drückb...*
    - TERM Tastatur
  - LANGUAGE FR
    - TERM clavier
  - LANGUAGE GA
    - TERM méarchlár
  - METADATA
    - IS PART OF EN computer FR ordinateur DE Computer Rechner GA ríomhaire CS počítač
    - IS KIND OF EN computer peripheral FR périphérique d'ordinateur

Figure 2.15: *The 'My Terminology Database' stock*

In lexicography, the usual way to organize entries is to start with a headword and then to describe all its meanings (called *senses*). In terminology, the approach is normally the opposite: you start with a single meaning (called a *concept*) and then you list off all the terms that express that concept, in all languages you are interested in. This is how the *My Terminology Database* stock is organized. The top-level element is a *concept* which contains several *language* sections. Notice that a *language* element contains a reference to a language name from the *language* category. Each *language* element contains one or more terms and, optionally, a definition (which is meant to be in the same language).

Terminologists are often interested in how concepts relate to one another: whether one concept is **part** of another concept (like a *processor* is part of a *computer*), whether one concept is a **kind** of another concept (like a *laptop* is a kind of *computer*), and many other relations. In our simple terminology stock, this is handled in the *metadata* section which can optionally appear at the end of the concept. This contains elements whose values are references to other concepts.


Firefox
My Terminology Database
lxln.prettydata.eu/Default.aspx
Global Glossary (English → Czech)
michal (Michal Boleslav Měchura) | log in again | log out | start page
My Terminology Database
rename
delete
CATEGORIES
new
concept
settings
language
settings
XML
save
delete

NAME	CARDINALITY		VALUE	PREVIEW?	APPEARANCE
ELEMENT CONCEPT			no value	yes	normal
ELEMENT LANGUAGE	required	repeatable	reference to language	yes	normal
ELEMENT DEFINITION	optional	one only	long text	no	*italic*
ELEMENT TERM	required	repeatable	text	yes	normal
ELEMENT METADATA	optional	one only	no value	no	normal
ELEMENT IS PART OF	optional	repeatable	reference to concept	no	normal
ELEMENT IS KIND OF	optional	repeatable	reference to concept	no	normal

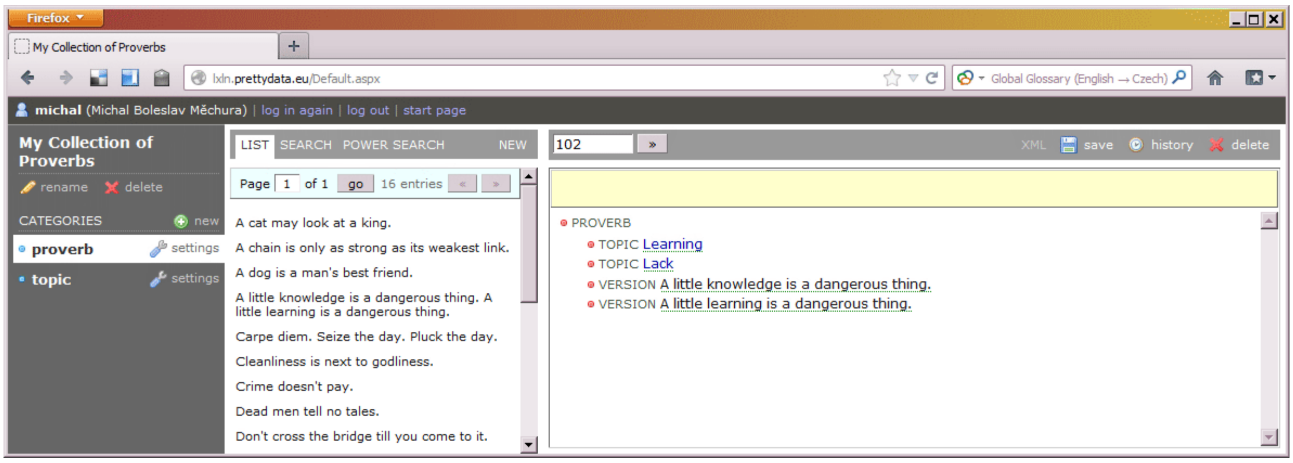
Figure 2.16: *Structure of the 'My Terminology Database' stock*

It would be easy to extend this structure to make it more sophisticated. For example, you could annotate the terms inside concepts with labels that express whether the term is an abbreviation or a full-form term, or you could add acceptability labels to terms (*recommended*, *deprecated* etc.). You could also add domain labels to concepts (*Computing*, *Biology*, *Mathematics* etc.).

2.6.2 My Collection of Proverbs

Lexicography and terminology are very conventional applications for software like Léacslann, but Léacslann can be used for any kind of data that requires an arbitrary but strict structure. The *My Collection of Proverbs* stock demonstrates this. Obviously enough, this stock is a collection of proverbs. Each *proverb* entry is organized in such a way that it can have one or more *versions* of the proverb (but there must be at least one) and can optionally be labelled with one or more *topics*. The topics are references to entries of the *topic* category.

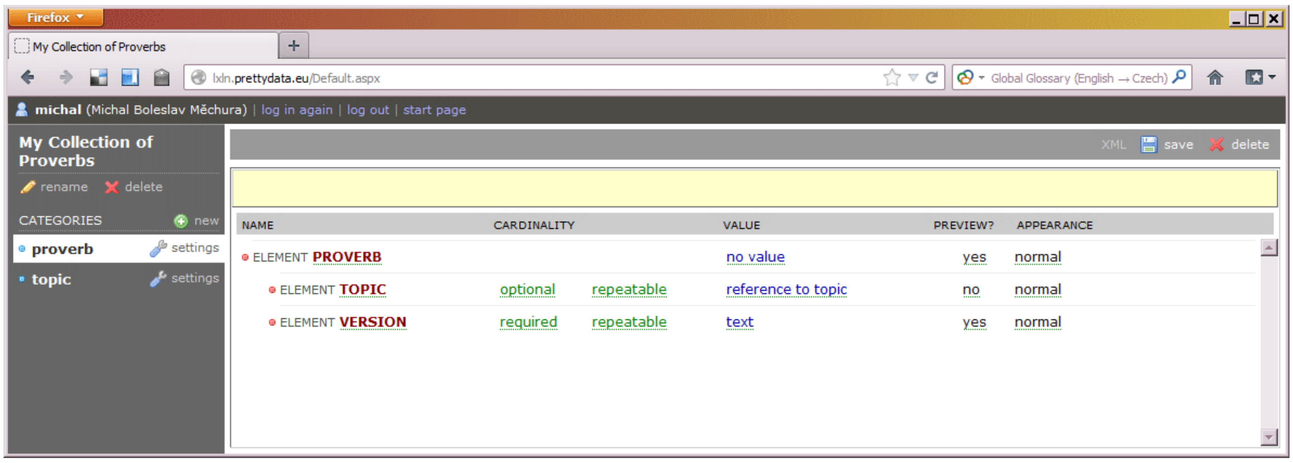
You could extend this to accommodate more data types. For example, if you are collecting proverbs from informants or from published sources, you could add an attribute to each version that tells you where that version came from – this would be a reference to entries of a category called *source* or some such. You could also label versions geographically or temporally, or you could add an element for comments. We will leave these suggestions for you to try out in your own time if you feel like it.


Firefox
My Collection of Proverbs
lxln.prettydata.eu/Default.aspx
Global Glossary (English → Czech)
michal (Michal Boleslav Měchura) | log in again | log out | start page
My Collection of Proverbs
rename
delete
CATEGORIES
new
proverb
settings
topic
settings
LIST
SEARCH
POWER SEARCH
NEW
Page 1 of 1 go 16 entries
102
XML
save
history
delete

A cat may look at a king.
A chain is only as strong as its weakest link.
A dog is a man's best friend.
A little knowledge is a dangerous thing. A little learning is a dangerous thing.
Carpe diem. Seize the day. Pluck the day.
Cleanliness is next to godliness.
Crime doesn't pay.
Dead men tell no tales.
Don't cross the bridge till you come to it.

- PROVERB
  - TOPIC Learning
  - TOPIC Lack
  - VERSION A little knowledge is a dangerous thing.
  - VERSION A little learning is a dangerous thing.

Figure 2.17: *The 'My Collection of Proverbs' stock*


Firefox
My Collection of Proverbs
lxln.prettydata.eu/Default.aspx
Global Glossary (English → Czech)
michal (Michal Boleslav Měchura) | log in again | log out | start page
My Collection of Proverbs
rename
delete
CATEGORIES
new
proverb
settings
topic
settings
XML
save
delete

NAME	CARDINALITY		VALUE	PREVIEW?	APPEARANCE
ELEMENT PROVERB			no value	yes	normal
ELEMENT TOPIC	optional	repeatable	reference to topic	no	normal
ELEMENT VERSION	required	repeatable	text	yes	normal

Figure 2.18: *Structure of the 'My Collection of Proverbs' stock*

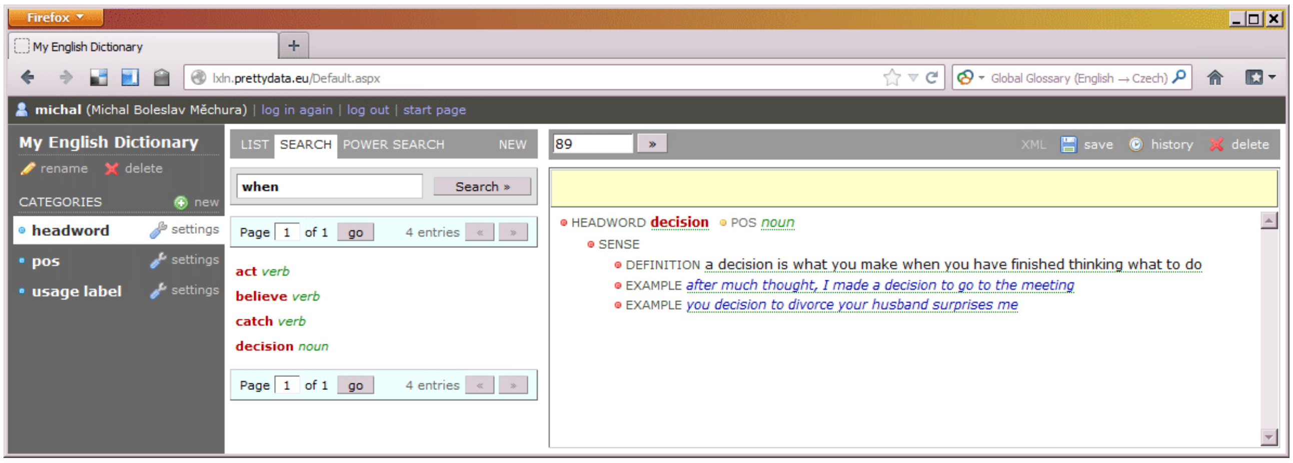
2.6.3 Over to you!

As you have no doubt figured out by now, you can create your own stocks in Léacslann by clicking the *Create a new stock* link on the start page. This will create a new stock which you must then populate by creating one or more categories and entries. The stocks you create are yours only and cannot be seen by other people, so don't be afraid to experiment.

3 SEARCHING AND LISTING ENTRIES

As you know by now, Léacslann displays an alphabetical list of entries in the middle of the screen. Once you've built up a large collection of entries in a stock in Léacslann, you will probably find that it is no longer practicable to wade through a complete list of all entries, and you will be wanting to filter and search entries based on some criteria. Luckily, Léacslann offers powerful search features, and we will introduce them in this chapter. The search features can be accessed by clicking the *Search* and *Power Search* links at the top of the screen.

3.1 Fulltext search



when

act *verb*
believe *verb*
catch *verb*
decision *noun*

HEADWORD **decision** POS *noun*
SENSE
DEFINITION a decision is what you make when you have finished thinking what to do
EXAMPLE *after much thought, I made a decision to go to the meeting*
EXAMPLE *you decision to divorce your husband surprises me*

Figure 3.1: *Fulltext search*

The first one (*Search*) gives access to a basic fulltext search facility which simply searches all text in all entries (of the given category) for the occurrence of the word or words you have entered. For example, when you type the word "when", Léacslann will find all entries where the word "when" occurs, regardless of where in the entry it is.

If you type more than one word, Léacslann will find entries where the words occur within the same element or attribute, even if they are not next to each other. For example, when you search for "act perform", Léacslann will find the headword *act* because both these words occur in one of its definition.

Note that Léacslann searches for whole words, not for parts of words. A search for "when" will yield results but a search for "whe" will not. Also, note that Léacslann will not find inflected or similar forms of the word. *(We are planning to introduce more sophistication into the fulltext search feature in a future version of Léacslann, including the ability to search for arbitrary substrings and the ability to expand queries with inflected forms.)*

3.2 Power search

The power search feature is where the power and flexibility of Léacslann really comes to shine. It allows you to search for entries based on a more-or-less arbitrary combinations of criteria. You can search for entries that contain a given substring or value in a particular element or attribute, for entries that have or do not have an element that fulfils some criteria, or for entries that have or do not have a particular number of elements that fulfil some criteria.

You can even compose queries that follow the train of references, so you can search for entries that refer to entries that fulfil some criteria. What's more, a search query can contain an unlimited number of criteria, which means that the potential for composing queries in Léacslann is practically endless. We will introduce the power search feature with a series of examples. All of the examples in this subchapter will be happening in the *My English Dictionary* stock.

3.2.1 Searching with text

Let's say you want to find *headword* entries where the value of the *headword* element contains some text, such as headwords that end with the string "er" (*brother*, *cooker*, *mother* etc.). We'll take you there step by step.

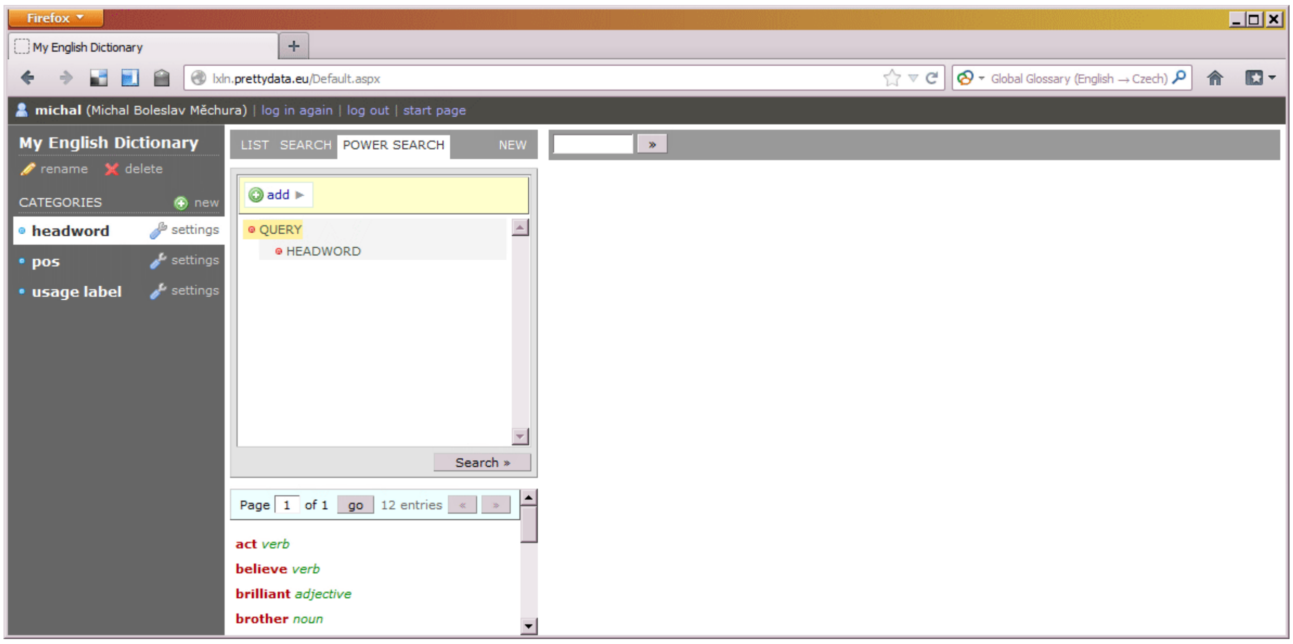
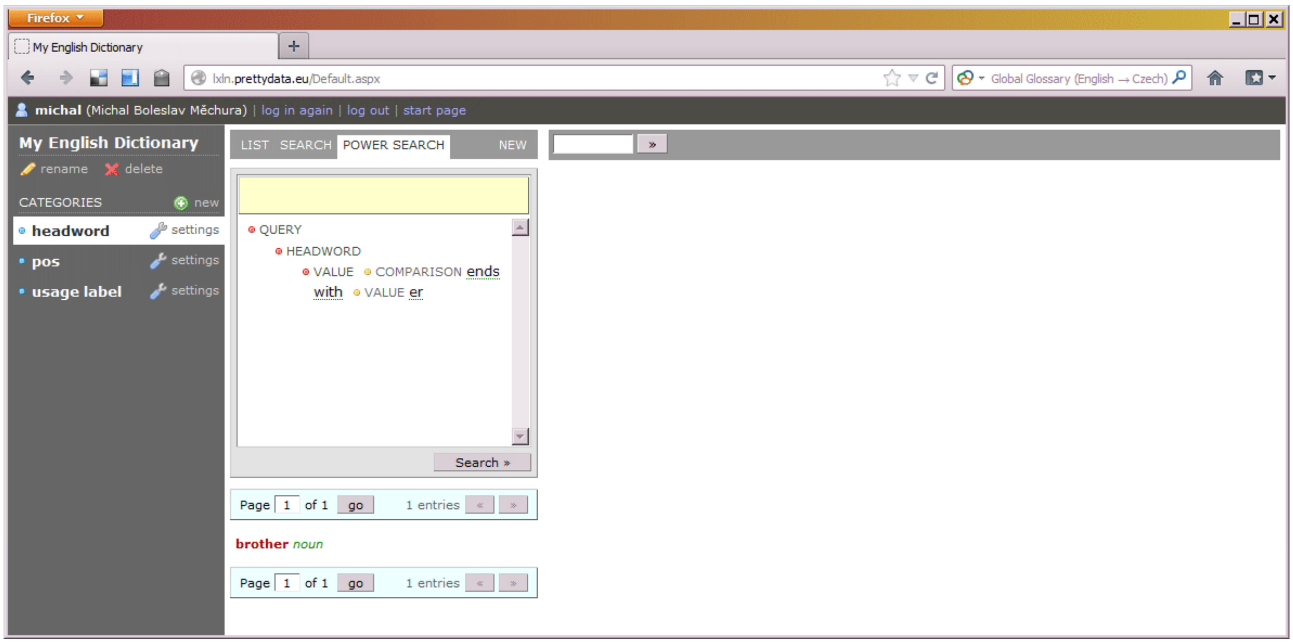


Figure 3.2: *Building a query, step 1*

Go to *Power Search* and highlight the *query* element. A button labelled *add* will appear in the toolbar at the top. Click it and then click *headword*. This has added a *headword* element under *query*. What you have done so far is create a query that says: "Find me every entry that contains a *headword* element!" If you want, you can click the *Search* button now to evaluate the query. Léacslann will return a list if all *headword* entries because, logically enough, all *headword* entries contain a *headword* element.

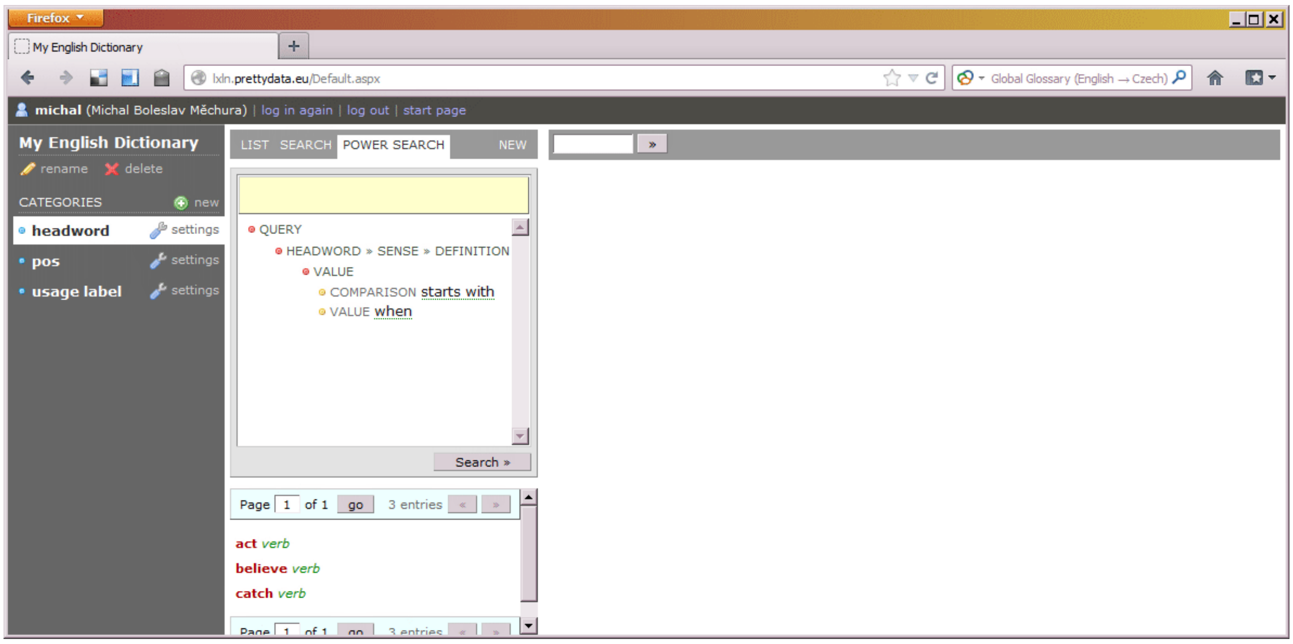
What you want to do now is to elaborate the query such that it says: “Find me every entry that contains a *headword* element whose value ends in ‘er’!” To do this, highlight the *headword* element, click the *add* button and then *value*. A new bullet point will be added underneath *headword*. Change *comparison* from *starts with* to *ends with* and replace the “click here” placeholder with “er”. You have now composed the query you wanted and you can evaluate it by clicking the *Search* button.



QUERY
- HEADWORD
  - VALUE
    - COMPARISON ends with
    - VALUE er

brother *noun*

Figure 3.3: Building a query, step 2



QUERY
- HEADWORD » SENSE » DEFINITION
  - VALUE
    - COMPARISON starts with
    - VALUE when

act *verb*
believe *verb*
catch *verb*

Figure 3.4: Finding definitions that start with ‘when’

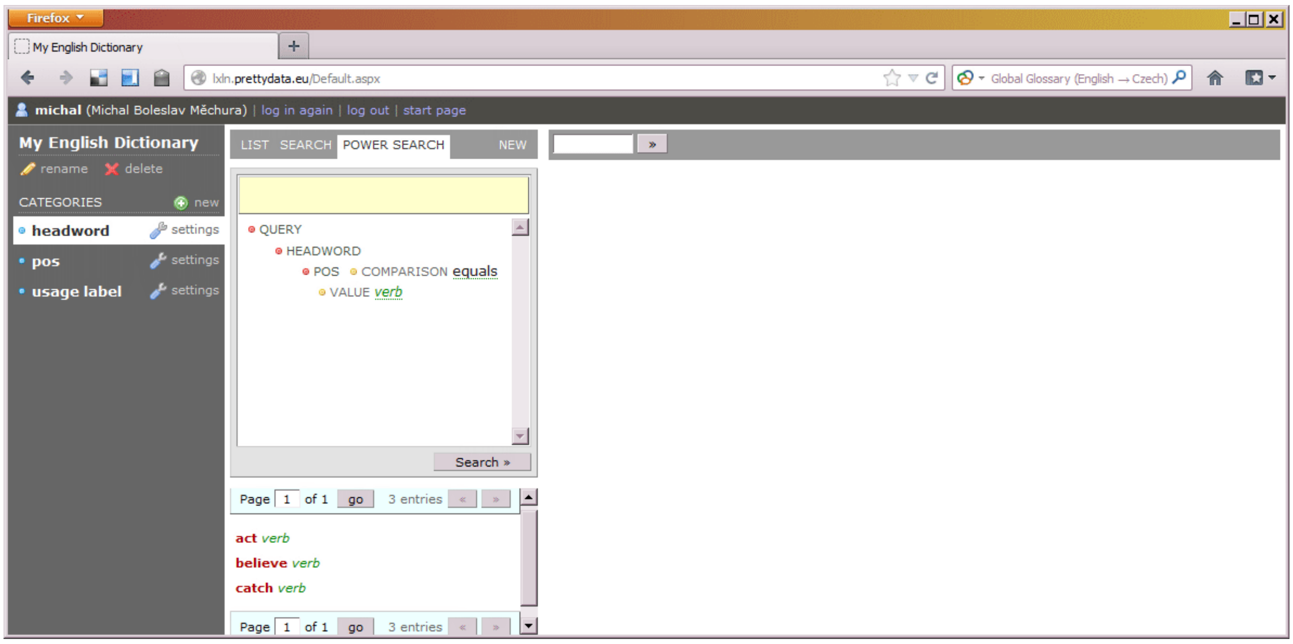
You can do searches like these on elements other than *headword* as well. Let’s say you want to find *headword* entries that contain a definition that begins with “when”. This is easy to do. Remove everything from underneath *query*, then highlight *query*, click the *add* button, then click *headword*

» *sense* » *definition*. You can figure out the rest from Figure 3.4. This query tells Léacslann: "Find me every entry that contains a *headword* element that contains a *sense* element that contains a *definition* element whose value starts with 'when'!"

3.2.2 Searching with values other than text

You are probably beginning to get the hang of it now. You compose queries by telling Léacslann to find entries that contains certain elements whose values fulfil certain criteria. For the sake of practice, we will look at a few more examples now, this time examples that involve values other than text.

Let's say you want to list all verbs. To do this, you want to compose a query that says: "Find me all entries that have a headword element that has, in its *pos* attribute, the value *verb*!" This is easy to do and Figure 3.5 will show you what the query is supposed to look like.

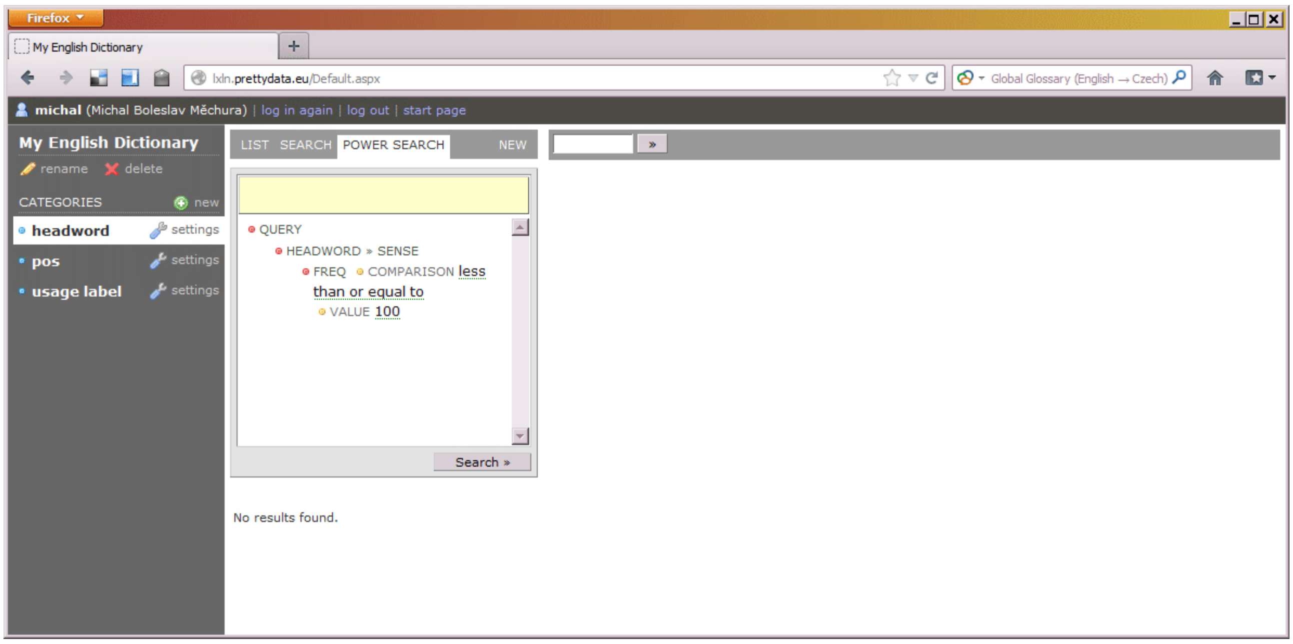


```
QUERY
  HEADWORD
    POS  COMPARISON equals
      VALUE verb
```

act *verb*
believe *verb*
catch *verb*

Figure 3.5: *Finding verbs*

Another query you might want to make is a query based on the *freq* attribute of *sense* elements, which tells you the corpus frequency of the sense (we created this in subchapter 2.3). Let's say you want to find entries where at least one sense has a frequency less than a given number. Figure 3.6 shows you what that query looks like. The query tells Léacslann: "Find me every entry that has a *headword* element that has a *sense* element that has, in its *freq* attribute, a value less than or equal to 100!"



QUERY
HEADWORD » SENSE
FREQ COMPARISON less than or equal to
VALUE 100

No results found.

Figure 3.6: *Finding entries by frequency*

Notice that, because Léacslann knows what data type each element and attribute has, it offers you the appropriate options in the appropriate place. When the data type is a number, Léacslann offers options (under *comparison*) such as *less than* and *greater than*. If the data type is text, the options are *begins with*, *ends with* and so on. This is another advantage to specifying the correct data type for each element and attribute (see subchapter 2.3).

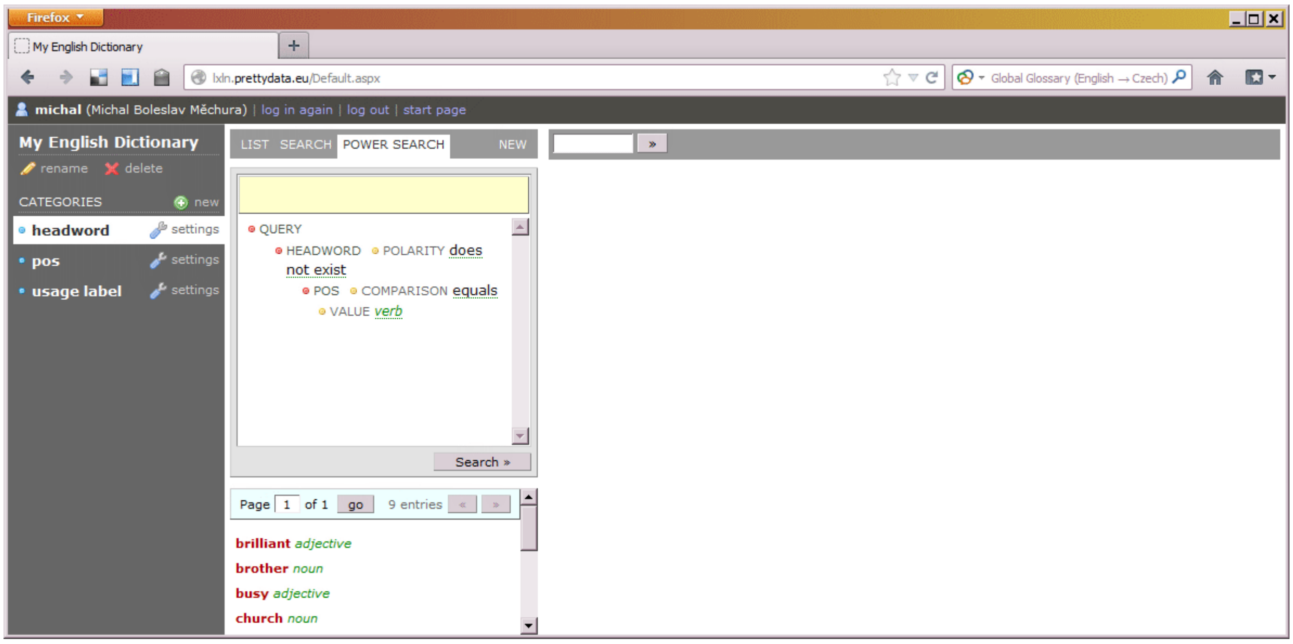
3.2.3 Negative searches

So far, you have been searching for entries that have elements that fulfil certain criteria. In addition to that, Léacslann allows you to make negative searches, that is, searches for entries that do **not** have an element that fulfils certain criteria.

We can show this on an example. Looking back at the query we built to find headwords that are verbs, we may want to find all headwords that are **not** verbs. The way to do this is as follows. Build the query as if you were looking for verbs, and then add an attribute called *polarity* to *headword* and change its value from *exists* to *does not exist*. This tells Léacslann: “Find me every entry where the following does **not** exist: a *headword* element that has, in its *pos* attribute, the value *verb*!” (Figure 3.7).

3.2.4 Searches based on number of elements

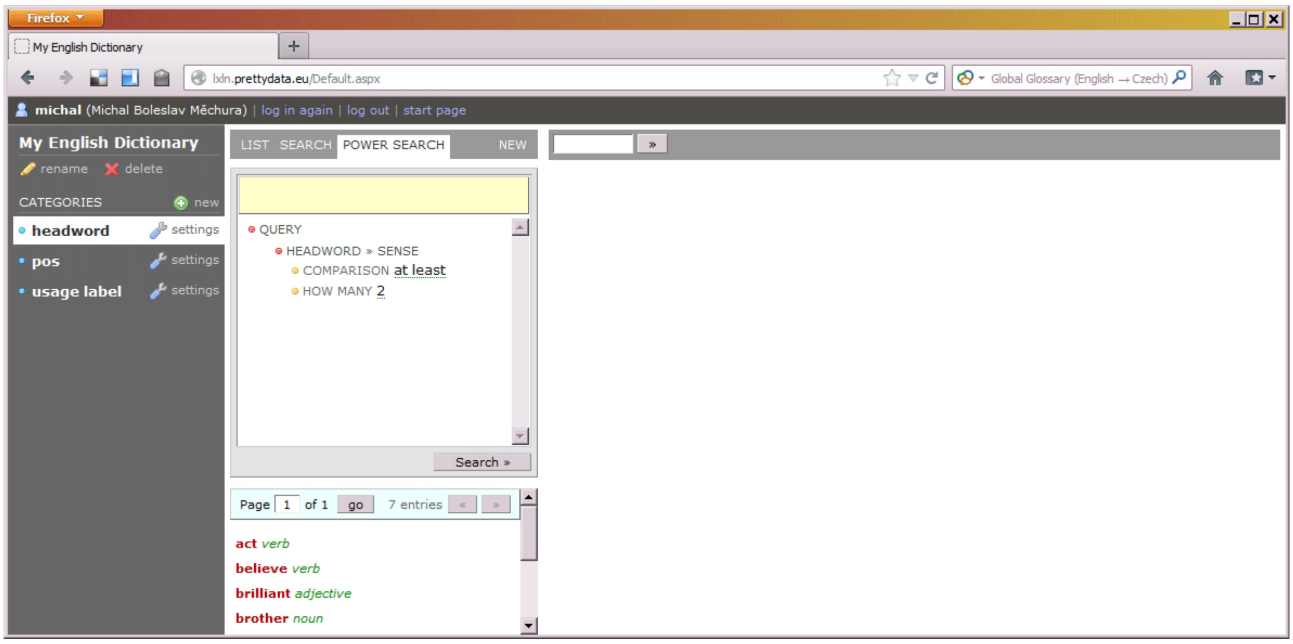
Another kind of search you can do in Léacslann is search based on the **number** of elements that fulfil certain criteria. Let’s say you want to find entries that have more than one sense. Figure 3.8 shows how to build such a query. What this tells Léacslann is this: “Find me every entry where at least two instances of the following exist: a *sense* element underneath a *headword* element!”.



QUERY
HEADWORD POLARITY does not exist
POS COMPARISON equals
VALUE verb

brilliant adjective
brother noun
busy adjective
church noun

Figure 3.7: Negative search



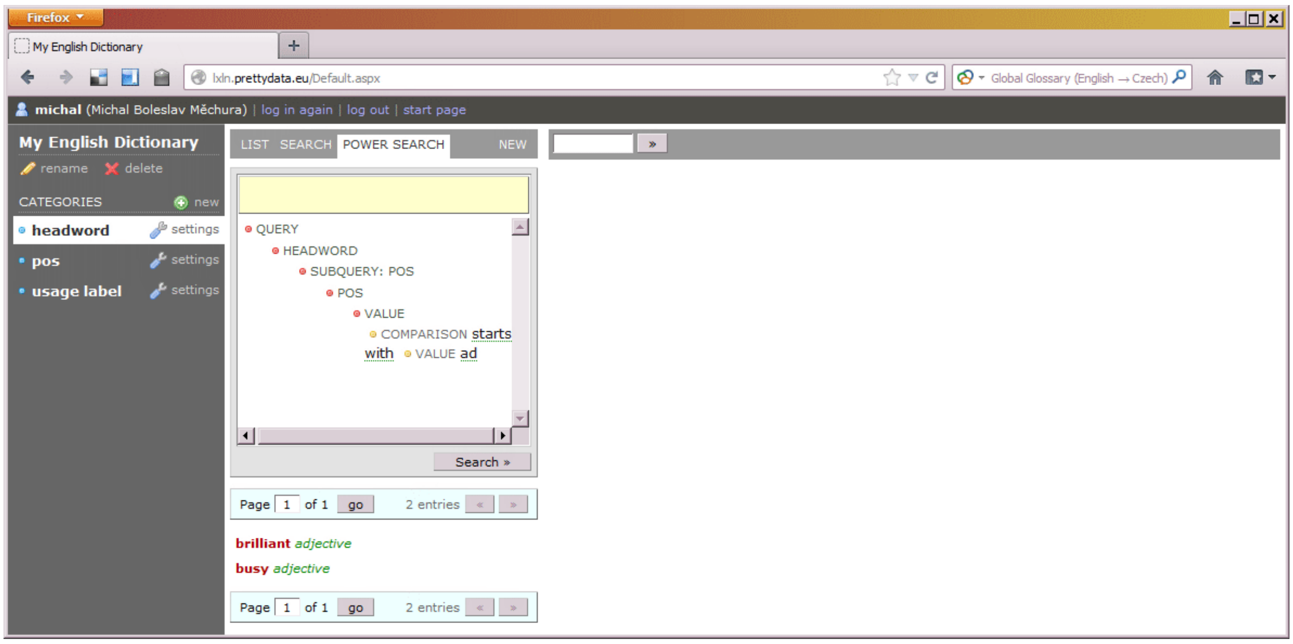
QUERY
HEADWORD » SENSE
COMPARISON at least
HOW MANY 2

act verb
believe verb
brilliant adjective
brother noun

Figure 3.8: Search based on number of elements

3.2.5 Searching with references

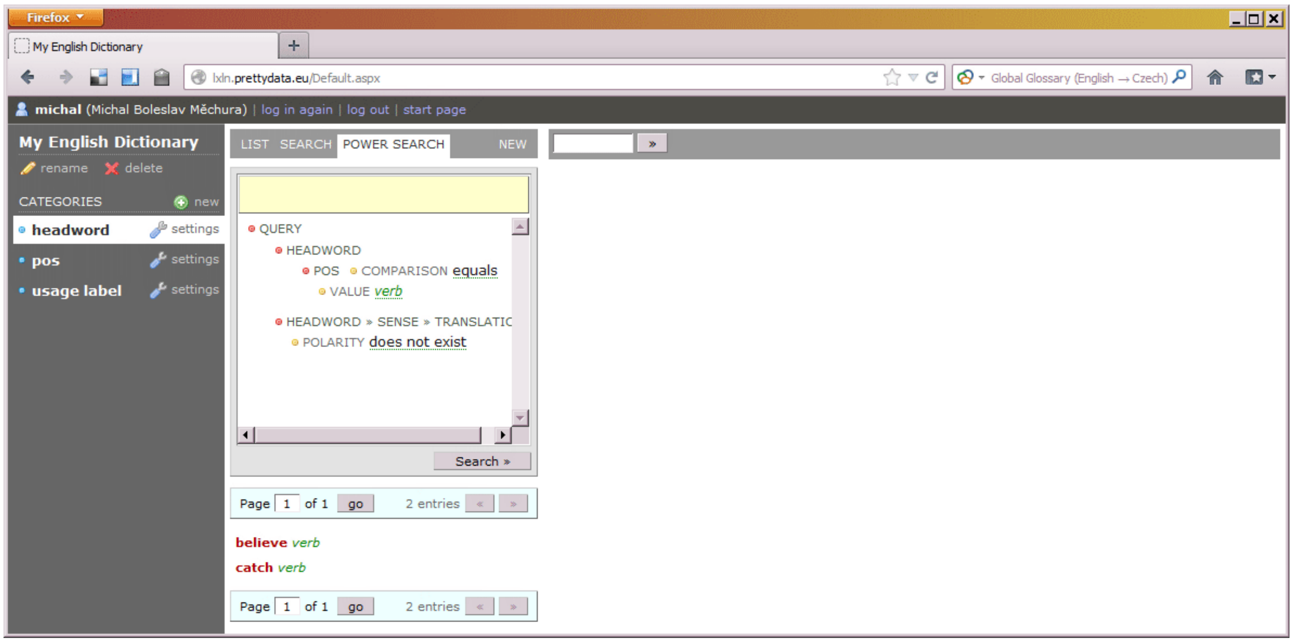
Léacslann's power search allows you to make searches that follow the train of references. What this means in practical terms is that you can search for entries that contain references to entries that fulfil certain criteria. We will demonstrate this on a slightly convoluted example: let's assume you want to find all headwords that have a part of speech whose title begins with "ad" (this includes both adjectives and adverbs). Figure 3.9 shows such a query. You will notice that this query contains something called a **subquery**. The whole query reads like this: "Find me all entries that have a *headword* element that has, in its *pos* attribute, a reference to a *pos* entry that fulfils the following criterion: it contains a *pos* element whose value starts with 'ad'!".



```
QUERY
    HEADWORD
        SUBQUERY: POS
            POS
                VALUE
                    COMPARISON starts
                with   VALUE ad
```

brilliant *adjective*
busy *adjective*

Figure 3.9: *Searching with references*



```
QUERY
    HEADWORD
        POS   COMPARISON equals
            VALUE verb

    HEADWORD » SENSE » TRANSLATIC
        POLARITY does not exist
```

believe *verb*
catch *verb*

Figure 3.10: *A query with two criteria*

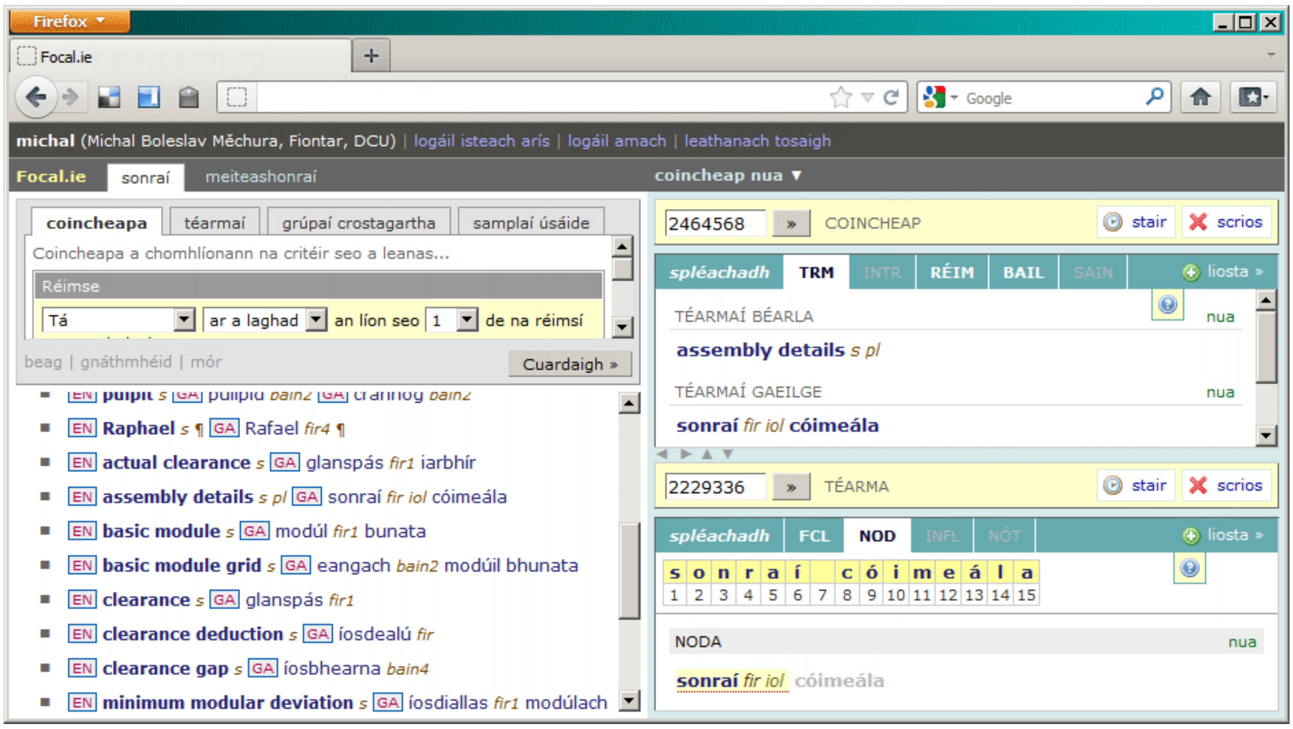
3.2.6 Combinations of criteria

When composing queries, you can combine multiple criteria in a single query. For example, if you want to find entries that are verbs and that have no translations, you can do it by building a query with two criteria: one that says "Find me every entry that has a headword element whose value, in the *pos* attribute, is *verb*!" and another one that says "Find me every entry where the following does not exist: a *headword* element that has a *sense* element that has a *translation* element!” Léacslann will return a list of entries that fulfil both criteria. Figure 3.10 shows what such a query looks like.

4 LÉACSLANN: THE BACK-STORY

What we have been showing you in this tutorial is only a small part of Léacslann. Léacslann is in fact a generic platform for building applications such as dictionary writing systems and terminology databases. The part of Léacslann which you have seen in this tutorial is only one such application, called the *Self-Service* application. It is possible to bypass the Self-Service application completely and to build your own customized application within Léacslann instead.

Deep inside, Léacslann is just a repository for storing structured entries, and a library of functions for accessing, changing and searching those entries in an application-neutral way. Léacslann was developed in Fiontar (an Irish-language teaching and research unit in Dublin City University) as a platform for working with multiple lexical databases, all of which are similar to, but slightly different from, one another. Fiontar mostly uses Léacslann as a platform for building highly customized applications; Figure 4.1 is a screenshot of one such application which we have developed to manage the National Terminology Database for Irish (available to the public at `http://www.focal.ie/`).



EN pulpit s GA pulpid bain2 GA crannóg bain2
EN Raphael s ¶ GA Rafael fir4 ¶
EN actual clearance s GA glanspás fir1 iarbhír
EN assembly details s pl GA sonraí fir iol cóimeála
EN basic module s GA modúl fir1 bunata
EN basic module grid s GA eangach bain2 modúil bhunata
EN clearance s GA glanspás fir1
EN clearance deduction s GA íosdealú fir
EN clearance gap s GA íosbhearna bain4
EN minimum modular deviation s GA íosdiallas fir1 modúlach

assembly details s pl

sonraí fir iol cóimeála

s	o	n	r	a	í		c	ó	i	m	e	á	l	a
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

sonraí fir iol cóimeála

Figure 4.1: *A highly customized Léacslann application*

Because Léacslann is a generic platform and because it can accommodate arbitrary data structures, we thought it would be a good idea to take Léacslann's potential to its logical conclusion and to build an application in Léacslann where users could to set up their own data structures in a user-friendly way, without the need to learn any programming languages or formal notations. That is the Self-Service application we have demonstrated in this tutorial.

The Self-Service application is experimental and still in development. You are very welcome to give it a test drive at the address `http://lxln.prettydata.eu/` and we will very much welcome your feedback. However, please be aware that it is likely to undergo changes in the future and that we cannot guarantee that your data will survive those changes intact. If you are considering using Léacslann for a real-world purpose (commercial or non-commercial) rather than for testing, please contact us to discuss the options.

4.1 Contact details

Michal Boleslav Měchura
E-mail: `mechrm@dcu.ie`
Postal address: Fiontar, Dublin City University, Dublin 9, Ireland