System IA-TAGGER; alfa version

User’s manual (29.12.2014)

Contents

[1. Roles of system users 3](#_Toc407121186)

[2. Getting stated 3](#_Toc407121187)

[3. How to upload a text document 3](#_Toc407121188)

[4. How to open a document 3](#_Toc407121189)

[5. How to save a document 3](#_Toc407121190)

[6. Document editing 3](#_Toc407121191)

[6.1. Division into sentences 3](#_Toc407121192)

[6.2. Editing contents of a sentence 4](#_Toc407121193)

[6.3. Word breaking 4](#_Toc407121194)

[6.4. Word annotation 4](#_Toc407121195)

[6.5. Sentence annotation 4](#_Toc407121196)

[7. Suggestions generated by the system 5](#_Toc407121197)

[8. System configuration 5](#_Toc407121198)

[8.1. Users 5](#_Toc407121199)

[8.2. Languages 5](#_Toc407121200)

[8.3. Word annotation levels 5](#_Toc407121201)

[8.4. Sentence annotation level 6](#_Toc407121202)

[9. Statistics 6](#_Toc407121203)

[9.1. Word statistics 6](#_Toc407121204)

[10. To be done 7](#_Toc407121205)

# User roles

Each user has a role ascribed. I can be *administrator*, *editor* or *reader*. Roles of users are ascribed only by the *administrartor*.

*Reader* – can only vew documentsCzytelnik może tylko przeglądać wszystkie dokumenty.

*Editor* – can edit the documents (chapter 6).

*Administrator* – has access to coiguration (Chapter 7).

# Getting started

System can be accessed from:

<http://rjawor.vm.wmi.amu.edu.pl/tagging/>

It requires: Chrome, Opera or Mozilla Firefox.

To start log in. The default passoword is tagger whcih can be changed at any time.

**login: <request login from system administrator>**

**password: tagger**

# How to upload a text document

Select element **Documents**. In the **File** window select a text document whihc you want to upload. In the **Language** window selects the language of the document. Click on **Submit**, to confirm the choice of the file and language. Selected file will be uploaded. The name of the file will be given at the end of the list of documents uploaded in the system.

During uploading the text is divided into sentences and words. One line is equal to one sentence and a sequence of signs between white signs is interpreted as a word.

# How to open a document

Click on the menu **Documents** and click on the naem of the document from the list (or the adjacent icon).

# How to save a document

You don’t have to save the document. Each modification is automatically saved in the system.

# Document editing

After opening the document is divided into sentences. Click on the icon of edition next to the sentence chosen by you. You goto the edition. Yoou can navigate between sentences usung arrows or using CTRL + arrow (up or down).

## Division into sentences

You can correct the actual division into sentences. Switch on the editing mode (by means of the icon *Edit mode* which is above the document and below the menu. Sentences can be divided by clicking on the scissors icon. Sentences can aslo be joined by clicking on the glue icon between sentences

## Editing contents of a sentence

You can change words in a sentence: add new ones or renove the existing ones.

Click on any word in a sentence. You will see the following icons:

– insert a new empty word before the current one (available for words which are not postpositions; 

add a word – available only for the last word

– remove (available only for words which do not have postpositions; to remove a word which has postposition first it has to be separated from the postposotion

 or (Ctrl-y) – mark the next word as a postposition (accessible only for words which do not have postpositions so far, or are not postpositions themselves, or are not at the end of the line

 or (Ctrl-u) – separate the word from the postposition (accessible only for words having postpositions or those which are postpositions).

## Word breaking

Words can be divided into stem and suffix: Click on the word and using mouse or keyboard slectt the place of the division and press Ctrl-J. Confirm with Enter. To remove the division press Ctrl-K and confirm with Enter.

## Word annotation

Words can be annotated at six different levels (levels can be set in *Configuration* menu, chapter 7).

To annotate a word at the selected level click on the level under the word you annotate.

Annotating at the level LEXEME = giving an English gloss

Annotating at the level of GRAMMAR – type the beginning of the tag and you will see the list of all tags beginning with the typed letters, e.g. writing the letter ‘f’ will cause displaying three tags: F (feminine), FOC (focus), FUT (future). Writting the sequence „pr” will cause displaying six tags beginning with „pr”. Click on the tag from the list to annotate the word.

Annotation at other levels consits in selecting one or more suggestions offered by the system. To select or deselect the offered suggestion click on it or press the shortcut button (given in the brackets).

Going to another field or pressing enter results in saving the value of the field in the database.

## Sentence annotation

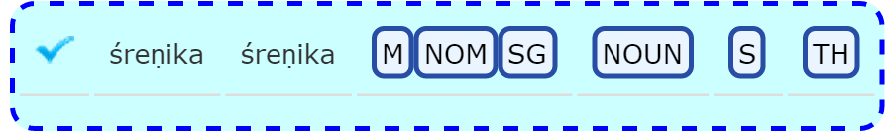
Snetences can be annotated at two levels (levels can be further set up in *Configuration* menu, chapter 7): *Add info*  nad *English* (English translation). At both levels sentence is annotated manually.

1. Tags suggestions generated by the system

IATagger system suggests tags’ values for words which fulfil one of the conditions:

1. Word was already tagged
2. Word’s structure allows the system to automatically deduce its values

Tags’ suggestions appear in a cloud above the word. For one word maximum three suggestions’ lines may appear in a cloud. Each line suggests values for a few tags, for example: for word śrenika one suggestion is:



You can accept the suggestion by clicking a ‘check’ symbol on the left side of a cloud which will assign following properties to tags:

LEXEME: śrenika  
GRAMMAR: N, NOM, SG  
POS: NOUN  
SYNTAX: S  
SEMANTICS:TH

1. System configuration

System configuration can be changed by an administrator. Other users don’t have access to configuration options.

By clicking **Configuration** you can open management menu with following parameters:

* 1. Users

This option allows to manage system users.

You can add a new user (+ *Add user*) or delete a user (*Delete*). You can ascribe or change a role in the system to each user (*Change Role*). You can reset user’s system access password to a default password (*Reset password*), for example when user forgets the password.

* 1. Languages

This option allows to manage languages of tagged documents.

You can add a new language (*+ Add language*). This language will be added to a list of languages supported by the system. Once a new language is added, opening and tagging documents saved in this language will be possible.

You can delete a chosen language (Delete) or edit it (Edit). Language edition may allow you to change a language code (Code) which helps in shorter identification of the language or to change the language description (Description)

## Word annotation levels

This option allows you to manage the levels of words’ annotation.

For each level you define:

* Name
* Description
* Strict Choice- if this option is selected (+), each word should have exactly one tag ascribed to this level
* Multiple Choice- if this option is selected (+), each word can have one or more tag ascribed to this level

Selection of the two options (*Strict choice* and *Multiple Choice*), even though technically possible, has no logical meaning. Not selecting any of these options means that ascribing a tag on this level is not compulsory.

Levels are placed in a particular order. For example, by default the POS level is set as the main/highest. You can change the levels’ order by using arrows in *Order* column.

You can add a new level (*+ Add word annotation level*) or delete an existing level (Delete).

You can edit level as well (Edit). By editing you can change name or description of a level or change options *Strict Choice* and *Multiple Choice.*

For each level you can add new or delete existing tags. To do so, you can use the option *Edit tags*. For each tag you should give its value (*Value*) and description (*Description*). In the process of tagging user will annotate words with tags’ values (*Values*) defined for each particular level.

You can change the order in which tags are suggested to user. For example, on the POS level’s list the first suggestion is NOUN tag. You can move it to the bottom of the list by using arrows.

## Sentence annotation level

This option allows you to manage levels of the sentence annotation, that is levels on which the Editor annotates whole sentences.

Managing sentence annotation levels is similar to managing word annotation levels. The only difference is that in the sentence annotation tags are not used.

# Statistics

You can generate statistics of selected/tagged texts by selecting option ‘Statistics’ from the main menu. There are two types of statistics: word statistics and colocation statistics.

## Word statistics

You can display word statistics by selecting a caption starting with a black dot, e.g.

* [**Verb participles (PTCP on the level "Grammar", V on "SYNTAX")**](http://rjawor.vm.wmi.amu.edu.pl/tagging/statistics)

System will give information about words number for which GRAMMAR tags’ value is *PTCP* and SYNTAX tag’s value is *V*. Moreover, it will display a list of words meeting the criteria along with values of all tags.

## Colocation statistics

You can display colocation statistics by selecting a caption starting with an empty dot, ex.

* [**PTCP(PTCP) + A(INS)**](http://rjawor.vm.wmi.amu.edu.pl/tagging/statistics)

System will give information about words’ pairs occurring in the same sentence, from which the first meets the criteria *PTCP(PTCP*) (values of both tags GRAMMAR and POS are *PTCP*) and the second meets the criteria *A(INS)* (value of SYNTAX tag is *A* and value of GRAMMAR tag is *INS*). Moreover, for each pair the system gives information about a distance between both words in the text (how many words are between them).