

## Reverse Transcription Tool Description

For each of the control buttons on the Reversible Transcription Intake page (e.g., Ottoman Arabic), there is a php script that is invoked when the button is clicked. The scripts are passed the contents of the Input Text Area and the contents of the Metadata fields and return the formatted output. These script files are as follows:

```
reverse_ottoman_arabic.php
reverse_ottoman_latin.php
reverse_ottoman_turkish.php
reverse_ottoman_cyrillic.php
reverse_xml.php
```

Each of these scripts works in a very similar fashion. Each loads a file of substitution rules that are used to perform the transcription. Any tags (e.g., "<italic>") in the input text area are removed and saved for later reinsertion. This is also true of html entity references (e.g., "&#1576;"). Both of these are removed so the application of the substitution rules will not mangle them. Incidental whitespace in the input is simply discarded. The substitution rules are then applied. The rules are in the form:

```
"B7" -> "+1576;", /* Arabic Letter Beh */
```

In this example, a "B7" in the input will be replaced by "+1576;" which is an entity reference for the Arabic Letter Beh. The "+" will be replaced by "&#" right before the final output. "\*" and "\*\*\*" are used to mark line breaks and line group breaks respectively. Line breaks are done with <div> tags and line groups have a <p> and </p> surrounding the groups of lines, each with its own <div> tag. The code in each script is commented if more detail is needed about its specific operation.

In the file of substitution rules, the rules are listed in the order they will be applied, separated by commas. Anything between "/" and "/" is a comment and can be used to add documentation. Each part of the rule is in quotes so that white space in the input can be interpreted correctly as needed. These are two distinct substitution rules:

```
" B7" -> "+1576;",
"B7" -> "+1576;",
```

The rules are applied in the order they appear in the rules file. When adding rules, one must consider where to order them in the list. In the two rules shown above, applying the rule without the space first would almost certainly prevent the one with the space from ever being applied.

The rules files are as follows:

```
ottoman_arabic_rules.txt
ottoman_cyrillic_rules.txt
ottoman_latin_rules.txt
ottoman_turkish_rules.txt
```

The XML php script uses the same rules file as the Ottoman Arabic. There is also a set of rules in the file "dynamic\_chart\_rules.txt". This is used to dynamically generate a table of all the substitution rules with a brief descriptive note. This chart is intended primarily for debugging purposes for the person modifying or adding rules to the transcription rules files.

## Functional Files

A complete list of the functional files (with a descriptive note) is as follows:

*reverse\_tool\_intake.html*

The intake page where text is entered to get a reverse transcription.

## PHP scripts:

*reverse\_ottoman\_arabic.php*

Script to return Ottoman Arabic.

*reverse\_ottoman\_latin.php*

Script to return Ottoman Latin.

*reverse\_ottoman\_turkish.php*

Script to return Turkish.

*reverse\_ottoman\_cyrillic.php*

Script to return the Cyrillic output. This currently in development.

*reverse\_xml.php*

Script to return XML. This currently in development.

*functions.php*

Has a couple of auxiliary functions.

## Substitution Rules text files:

*ottoman\_arabic\_rules.txt*

*ottoman\_cyrillic\_rules.txt*

*ottoman\_latin\_rules.txt*

*ottoman\_turkish\_rules.txt*

*dynamic\_chart\_rules.txt*

## Minimal stylesheets:

*style\_arabic.css*

*style\_cyrillic.css*

*style\_latin.css*

*style\_turkish.css*

## Site Map

**<http://courses.washington.edu/otap/reverse/reverse.html>**

First page of the Reverse Transcription Tool featuring a project description and a link to an article about reverse transcription.

**[http://courses.washington.edu/otap/reverse/reverse/o\\_Reverse\\_trans\\_article728.html](http://courses.washington.edu/otap/reverse/reverse/o_Reverse_trans_article728.html)**

The above mentioned article.

**[http://courses.washington.edu/otap/reverse/intake/reverse\\_tool.html](http://courses.washington.edu/otap/reverse/intake/reverse_tool.html)**

Page with the following links for more information and link to the tool itself.

**[http://courses.washington.edu/otap/reverse/intake/getting\\_started.html](http://courses.washington.edu/otap/reverse/intake/getting_started.html)**

Instructions for using the reverse transcription tool.

**[http://courses.washington.edu/otap/reverse/intake/consonants\\_eng.html](http://courses.washington.edu/otap/reverse/intake/consonants_eng.html)**

Chart of consonant conversions in English order.

**[http://courses.washington.edu/otap/reverse/intake/consonants\\_arb.html](http://courses.washington.edu/otap/reverse/intake/consonants_arb.html)**

Chart of consonant conversions in Arabic order.

**<http://courses.washington.edu/otap/reverse/intake/vowels.html>**

Chart of vowel conversion codes.

**[http://courses.washington.edu/otap/reverse/intake/special\\_characters.html](http://courses.washington.edu/otap/reverse/intake/special_characters.html)**

Chart of special character codes.

**[http://courses.washington.edu/otap/reverse/intake/conversion\\_chart.php](http://courses.washington.edu/otap/reverse/intake/conversion_chart.php)**

Dynamically generated chart that displays all of the substitution rules.

**[http://courses.washington.edu/otap/reverse/intake/writing\\_arabic.html](http://courses.washington.edu/otap/reverse/intake/writing_arabic.html)**

Incomplete article.

**[http://courses.washington.edu/otap/reverse/intake/reverse\\_tool\\_intake.html](http://courses.washington.edu/otap/reverse/intake/reverse_tool_intake.html)**

The intake page where text is entered to get a reverse transcription.

**[http://courses.washington.edu/otap/reverse/intake/reverse\\_ottoman\\_arabic.php](http://courses.washington.edu/otap/reverse/intake/reverse_ottoman_arabic.php)**

Script to return Ottoman Arabic. Run this by clicking on the Ottoman Arabic button.

**[http://courses.washington.edu/otap/reverse/intake/reverse\\_ottoman\\_latin.php](http://courses.washington.edu/otap/reverse/intake/reverse_ottoman_latin.php)**

Script to return Ottoman Latin. Run this by clicking on the Ottoman Latin button.

**[http://courses.washington.edu/otap/reverse/intake/reverse\\_ottoman\\_turkish.php](http://courses.washington.edu/otap/reverse/intake/reverse_ottoman_turkish.php)**

Script to return Turkish. Run this by clicking on the Turkish button.

**[http://courses.washington.edu/otap/reverse/intake/reverse\\_ottoman\\_cyrillic.php](http://courses.washington.edu/otap/reverse/intake/reverse_ottoman_cyrillic.php)**

Script to return the Other output. This currently in development.

**[http://courses.washington.edu/otap/reverse/intake/reverse\\_xml.php](http://courses.washington.edu/otap/reverse/intake/reverse_xml.php)**

Script to return XML. This currently in development.