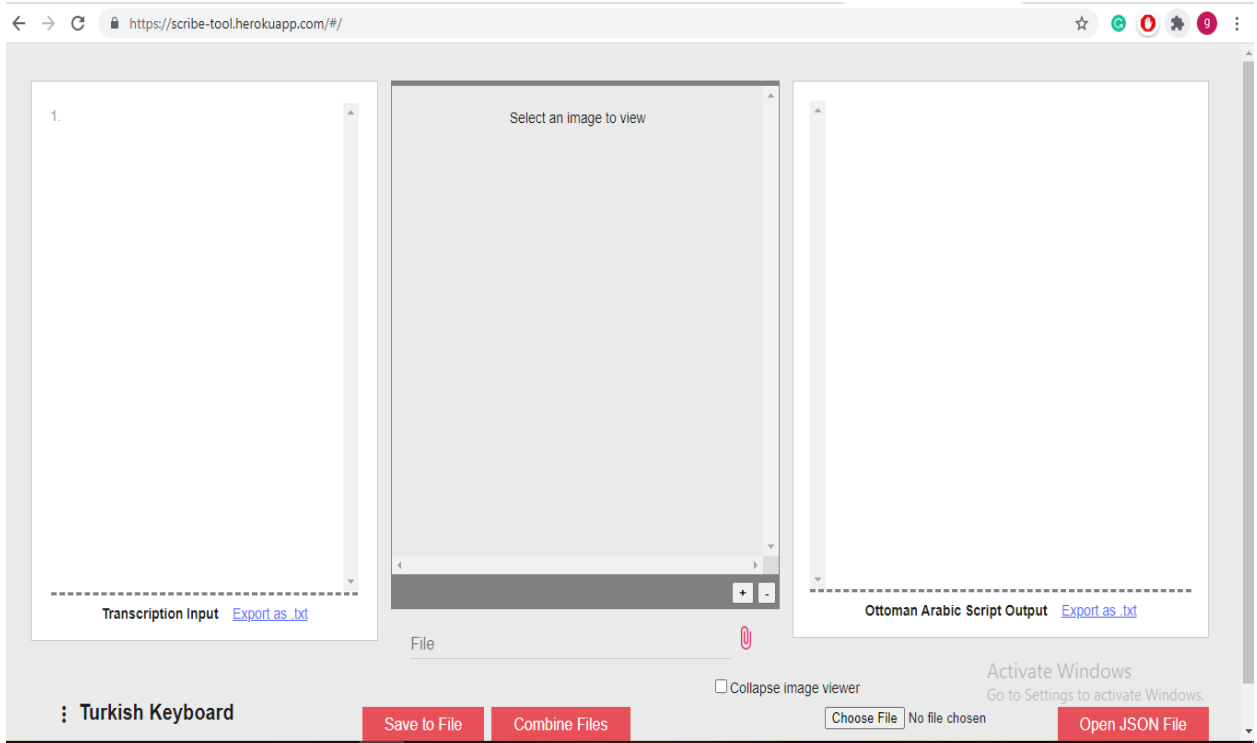


GETTING STARTED WITH *SCRIBE* [TURKISH Q KEYBOARD]

SCRIBE BETA TEST VERSION

<https://scribe-tool.herokuapp.com>



[NOTE ON THE BETA TEST VERSION: WE WILL STILL BE MAKING SOME CHANGES IN THE TOOL.]

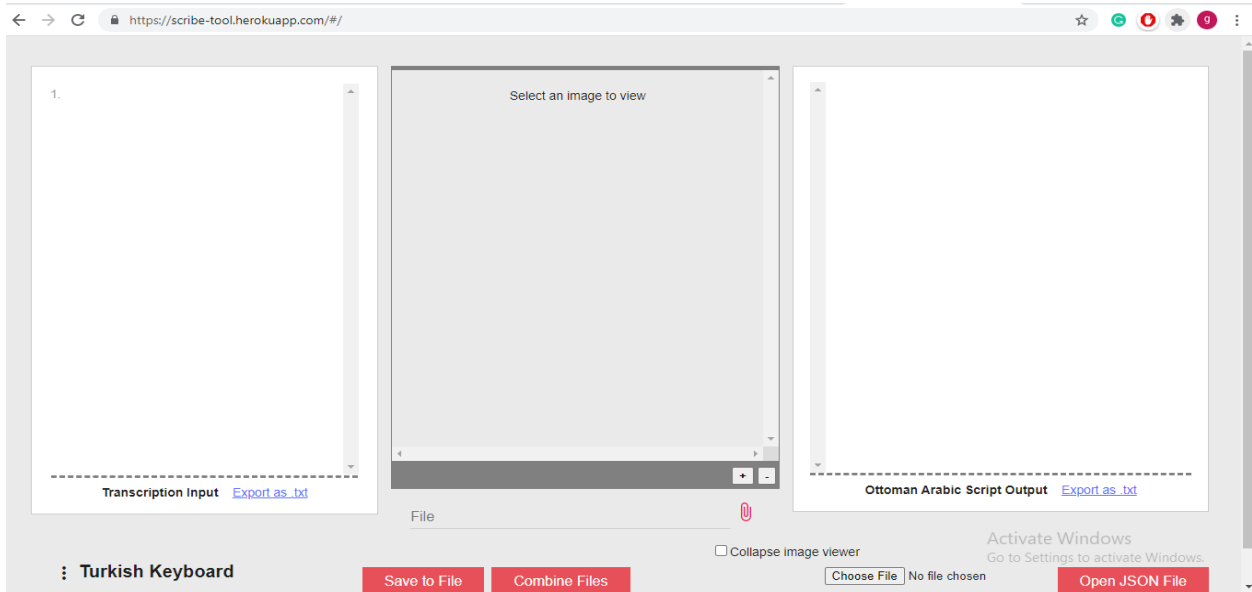
GETTING STARTED WITH *SCRIBE* [TURKISH Q KEYBOARD]

INTRODUCTION

SCRIBE is an application intended to improve the accuracy of transcription from Ottoman Turkish Arabic script manuscripts to Latin-based fonts by linking the Latin transcription directly to the Arabic characters it is transcribing. This means that as a researcher types Latin characters in one text box, the corresponding Arabic characters appear in another text box. Images of the original manuscript text can also be displayed by the tool allowing a direct comparison between the Arabic script original and the Arabic script transcription created by the tool. The Latin and Arabic versions can be exported as plain text and the core data can be returned to the tool for editing and correction. The result is digital transcriptions in both the Arabic script and a version of the Turkish transcription alphabet.

SCRIBE uses selections from “dropdown menus” to produce the Latin letter input. This means that any standard font will produce the required Unicode (UTF8) Latin and Arabic script text. One does not need a special font on one’s device. The tool is now configured to work best with the standard Turkish Q keyboard. The dropdown scheme is borrowed from applications developed to deal with the complexity of Chinese characters. *(The choices may seem many and confusing at first but remember, most of the choices you will make are limited and you will soon learn to enter them quickly. There are many choices necessary only for entering Arabic text, which makes the system look more complicated than it really is).*

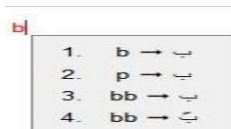
The best way for you to become acquainted with *SCRIBE* is to begin by playing with the tool. It is much easier to learn by doing than to try to understand complex directions. So, let’s start by looking at the tool:



GETTING STARTED WITH *SCRIBE* [TURKISH Q KEYBOARD]

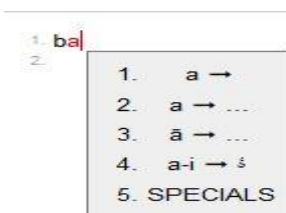
To get started you simply start to type. Let's begin by entering the name of the poet whose name in the original script looks like this باقى (Bākī).

You will begin by typing the letter “b” and a “dropdown menu” will appear looking like this:



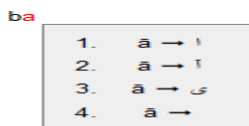
The first rule you need to know is that the #1 option in every dropdown is the most common one and you do not need to enter the 1. Just go on to a space or to the next letter. Note the other options you have when you type “b”. You can select a “p” when you want to reflect the pronunciation of a “ba” when it is final or unvoiced in Turkish words (for example, “olup”). Or you can select cases where the “ba” is doubled by “shedde”.

When you type the “a” a “ba” will appear on the Ottoman Arabic [O-A] script screen on the far right and you will get the following dropdown menu from which you can select what kind of an “a” you want. Note that the #2 option produces “vocalic he” when it is in a Turkish word. And does a22 produces an “elif” for Turkish words that does not show as “long” in the transcription.



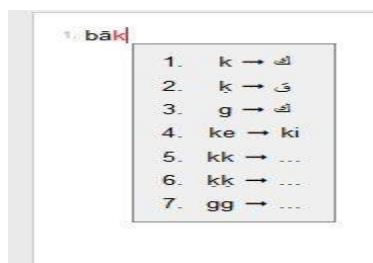
If you just continue you will get a “short a” which is not represented by a character in the O-A script. But here you want a “long a” represented by “elif” in the O-A script so you would type a3 after the “a” and continue. You will see “ba” + “elif” “با” on the O-A screen. *(Before you continue, however, notice the cases where there are three dots [...] after the first choice. The three dots [...] mean that there are more choices than can fit in a box (we try to limit the boxes to 9 choices). For example, you will get another menu with more choices when you choose “a3”.)*

If you choose #2 you will get an “elif” with “medd” (آ). If you choose #3 you will get an “elif maksure” (ى).

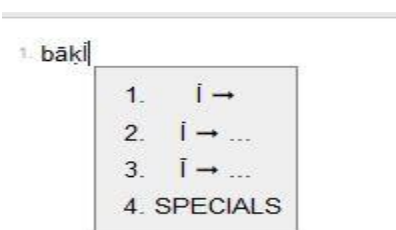


GETTING STARTED WITH *SCRIBE* [TURKISH Q KEYBOARD]

Because you want the “long a” (ā) you will just continue by typing a “k”, which will produce the following choices:



Here there are many choices because of the different uses of the letter “k”. In this case, because you want to represent “kaf”, you would type “k2” and continue to “i”. Note that one of the options (#4) in this menu is for producing the particle “ki” (as in “dedi ki...”) which is written “kef” + “güzel he” (كف). When you type “i” you will get the following menu:



The options here are similar to those for “a”. You will want to type “i3” to get the long final “ī”.

Now it is time for you to go ahead and try some more transcribing on your own!

TRANSCRIBING TURKISH

The largest challenge for experienced transcribers will come in the way Turkish words are transcribed. At this point there is a major departure from common practice. The theory is quite straightforward and simple: The practice of transcribing Turkish vocabulary as it is *pronounced* without regard to the Arabic script letters used results in a significant loss of data about the historical orthography of Ottoman Turkish. A reversible transcription system like *SCRIBE* demands that the Arabic script transcription must represent exactly what is in the original text and this means that “long” (visible) vowels in the Arabic script versions of Turkish vocabulary must remain in the Arabic script transcription—although not in the Latin script versions. The same is true of “voiced” consonants that in Turkish pronunciation lose their voicing and become unvoiced. *SCRIBE* allows for retaining “long” vowels and “voiced” consonants in the Arabic script versions of Turkish words while keeping the vowels short and letting final or stopped voiced consonants

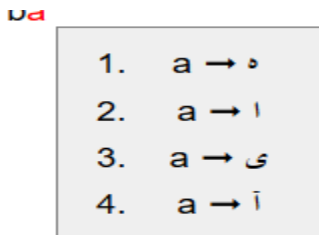
GETTING STARTED WITH *SCRIBE* [TURKISH Q KEYBOARD]

become unvoiced in the Latin script versions. **This will be different than usual transcription practice for Ottoman Turkish because you must type a special Arabic script letter code for any letters that appear in Turkish words in your text. This means that visible letters appearing in Turkish words, for example “ا”, “و”, “ى” and “ه”, must be entered.** (Eventually this will allow statistical studies of spelling and the history of spelling in Ottoman Turkish).

For experienced transcribers, this will, at first, be substantially more complex than their normal practice. But it is not that difficult either and the relevant codes are easily memorized. Let’s look at a very simple example. Take the phrase “Bākī baḳtī” and imagine that “baḳtī” is actually written “bāḳdī” (باقدى). You already know how to write “Bākī” [Ba3k2i3] but writing “baḳtī” in a way that preserves the spelling of the original Arabic script text involves another set of characters representing Arabic script characters as they are used in Turkish words. For example, an elif used as a vowel in a Turkish word is not “long” because Turkish words do not have “long vowels”. This elif is an indicator of a Turkish phoneme. Go back to the “a” dropdown.



You will see (#2) a “short a” followed by the three dots that indicate other choices. If you type “a2”, you will get the following dropdown:



The #2 option in this dropdown is an elif that we might call “Turkish elif” which is an elif [a22] that represents the Turkish phoneme “a” and has no “length” associated with it.

In the same way, the “d” is a “Turkish d” which is pronounced as “t” when “closed, final, or in an unvoiced environment”. You will see this option [#3] in the “d” dropdown. The final “Turkish “ı” is created in the same way as the “Turkish elif”. On the “ı” dropdown select #2 and continue. If you have typed ba22k2d3ı2, you should see **baḳtī** in the Latin transcription and باقدى in the Ottoman Arabic.

As you type transcriptions of Turkish vocabulary, you will, at first, need to pay close attention to selecting a Turkish version where one is necessary.

GETTING STARTED WITH *SCRIBE* [TURKISH Q KEYBOARD]

TRANSCRIBING PERSIAN

Because Ottoman Turkish regularly employs a wide variety of Persian elements it would be unwieldy and unnecessary to produce examples of the whole range of Persian grammar. There are a few points, however, that will help you get started:

The Persian Izafet (-i, -ı, (ī- /i), -yi, -yı, and with hemze)

(-ı/-i):

üstād-ı sūḥān (استاد سخن): ü4sta31d-i3 (space) sūh3an
‘īnāyet-i ḥudā (عنایت خدا): ‘ina31yet-i2 (space) h3uda31
pāy-ı endīše (پای اندیشه): pa31y-i3 (space) e22ndi31še21
ḥāne-i peder (خانه پدر): (without hamza): h3a31ne21-i2 (space) peder
ḥāne-i peder (خانه پدر): (with hamza): h3a31ne4 (space) peder

(-yı/-yi):

nevā-yı bülbül (نواى بلبل): neva31-yı (space) bülbül
āhū-yı çeşm (آهوى چشم): a32hu4-yı (space) çeşm

(ī- /i):

bākī-i ḥaste-dil (باقى خسته دل): ba31k2i31-i2 (space) h3aste21-dil
keşti-i nūḥ (کشتى نوح): keşti31-i2 (space) nu4h2
sākī-i meclis (ساقى مجلس): sa31k2i31-i2 (space) meclis
perişānī-i dil (پريشانى دل): peri31ša31ni31-i2 (space) dil

[To create this izafet with a hemze showing on the “long i” (“ī / i”), end the first element with i34-i2 as in keşti34-i2 nu4h2 => کشتى نوح]

The Archaic Persian Letter “خو ḥ^w“

Traditional Ottoman transcription has not fixed on a single solution for representing this letter. *SCRIBE* uses the transcription convention: ḥ^w (h4). Thus we have, for example:

h4i3ş => ḥ^wiş => خویش

h4a3b => ḥ^wāb => خواب

h4a3ce2 => ḥ^wāce => خواجه

GETTING STARTED WITH *SCRIBE* [TURKISH Q KEYBOARD]

TRANSCRIBING ARABIC

Transcribing Arabic and especially “vowelled” Arabic is challenging. You will need to become familiar with both the “w” menu and the apostrophe [‘] menu.

For experienced transcribers, working with *SCRIBE* will also require a few changes in the way you are accustomed to transcribing Arabic. The most obvious is in compounds with the definite article (al-). Let’s take a simple case which would ordinarily be transcribed “dārū’l-islām”. This accurately represents the *pronunciation* of the phrase but not its *orthography* (writing). *SCRIBE* always respects the written form of the text. In both cases, the “apostrophe” actually represents an *elif + vasle*, which, in the written form, is the first letter of a separate word. The phrase is written in the Arabic script like this: دارُ الإسلام. Go to *SCRIBE* and enter the following: da3rūw3 <space> ‘4-li45sla3m [= dārū 'l-islām]. Notice how the “w” dropdown gives you options for adding *harekat* (vowel marks) and how the Latin transcription follows the word-breaks in the original Arabic.

What about the definite article with the “sun” letters? Let’s try writing “dāru 'd-dünyā” (remember the space before the elif vasle): da3rū <space> ‘4d-d2ünya3. You should get **dārū 'd-dünyā** (Latin) and (Ottoman Arabic) دارُ الدُّنْيَا.

Now let’s try creating a fully vowelled version of the first misra of the famous Muallaqa of İmriü ‘l-Ḳays

K2i32fwa3 <space> ne52bw7ki43 <space> mi43nw7 <space> z4i43kw7rwa33 <space>
h2a5bw2i3bin6 <space> vwe <space> me52nw7zi43li43

ḳıfā nebki min zıkrā ḥabībin ve menzili

قَفَا نَبْكَ مِنْ ذِكْرَى حَبِيبٍ وَ مَنْزِلِ

If you see how this works, you are on your way to accurately transcribing Arabic and using the “w” dropdown.

TRANSCRIBING HAMZA

Another feature of Arabic that *SCRIBE* approaches in a way that deviates from common usage is in the representation of hemze and its “seats” (kürsi). *SCRIBE* uses the phonetic symbol for a glottal stop (ʔ) instead of an apostrophe, because in a digital environment it is better if each transcription symbol has only one use. It also uses special input to create the kürsi. These are demonstrated in the list below:

On the kürsi of elif: “أ”

teʔlīf (تأليف): teʔ5Li3f

meʔmūr (مأمور): meʔ5mu4r

On the kürsi of vav: “و”

muʔahḥar (مؤخر): muʔ7ah71ar

GETTING STARTED WITH *SCRIBE* [TURKISH Q KEYBOARD]

şü'ün (شؤون): şü'7u4n

On the kürsi of ya: “ی ”

ṭā'ife (طائفه): t2a31'6ife21

mes'ele (مسئله): mes'6eLe21

[Remember, you do not need to type a “1”, just add the following letter or space.]

NOTE: One other feature of Arabic that *SCRIBE* transcribes in an unusual way is the letter “tha” ث pronounced “s” in Turkish and usually represented in Ottoman transcription as an “s” with a bar under. Because the “s bar under” does not appear as a single letter in the common Latin Unicode fonts, we have chosen to use “s circumflex” “ŝ” to represent “tha”, as in “eŝer” for اثر . In the same way, instead of a right-facing apostrophe to represent “ayin” (ع), *SCRIBE* uses (‘) (called “right facing half ring”).

SAVING, REVISING, EXPORTING AS TEXT

The basic *SCRIBE* files are .json files. JSON (JavaScript Object Notation) is a data-interchange file format (see <http://json.org/> for details). The .json files contain the information necessary to display transcription text on *SCRIBE*. They are not directly readable by humans.

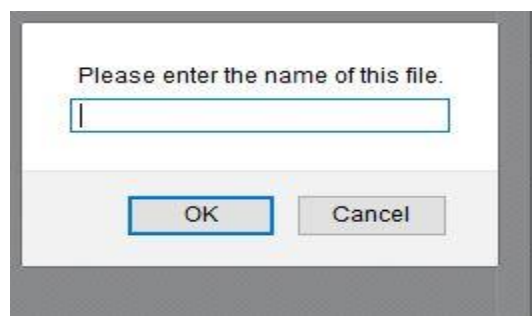
For readable/printable text output the .json files must be exported as .txt files or copy-pasted into a word processor from the *SCRIBE* display.

It is important to note that *the exported text files cannot be returned to SCRIBE for editing*. The .json files are the *only* editable *SCRIBE* source.

It is also important to note that *the exported files are no longer linked*. Any editing of the exported text files breaks the crucial link between the Latin and Ottoman Arabic transcriptions.

TO SAVE:

1. Simply start transcribing.
2. When you are ready to SAVE, click on the red button that says “Save to File”.
3. You will see the SAVE box.



GETTING STARTED WITH *SCRIBE* [TURKISH Q KEYBOARD]

4. Enter the name of your file and click “OK”. When you see the location menu, browse for a location and click SAVE.

TO EDIT THE FILE YOU CREATED OR RESUME TRANSCRIBING:

1. Go to the “Browse” button on the lower right of the *SCRIBE* screen. Click on it and browse to the location of your JSON file. Select and open the file. The filename will appear to the right of the “Browse” button.
2. Click on the red “Open JSON File” button on the lower right of the screen. Your file will appear in the two text boxes.
3. Edit the file or resume transcribing and when you are done go through the SAVE procedure again.

EXPORT YOUR FILES AS .TXT:

1. At the bottom right of each text pane [Latin and Arabic] is a live blue link that says “Export as .txt”. Click on the link.
2. You will get the same “filename box” that you got for saving your JSON file. Enter a name, for example, “filename_arabic” or “filename_latin” and click “OK”. ***Remember that the material in each text pane will export as a separate .txt file.***
3. Your .txt files can now be opened in any word processor. ***Remember that these files are no longer linked and cannot be returned to SCRIBE for further editing. You should consider your JSON file as the ultimate source for consistent and accurate transcription texts. Preserve your JSON files carefully.***

SELECTING AND LOADING AN ORIGINAL TEXT IMAGE:

1. The middle pane of the *SCRIBE* screen is for displaying images of the original (manuscript) texts being transcribed.
2. To select an image to view, go to the red “paperclip” image below the screen and click on it. You will get a navigation screen from which you can browse to the location of your file, which should be in a standard image file format: JPG, TFF, etc. We have not tested *SCRIBE* with all possible formats but you are always safe saving your files as JPG.
3. Click on your file and it will appear in the display pane. You can increase and decrease the size of the text by clicking on the “+” and “-“ symbols on the lower right of the pane.

[At this stage of development, SCRIBE does not have the capability of loading large image files (a whole divan for example) but we expect to add this in the near future.]

QUESTIONS, ANSWERS, HINTS, AND SUGGESTIONS