Dspeech User Manual

This program was designed by Dimitrios Coustoumbas

Written by Michel SAVARD

Information :

The installation of this program does not modify Windows. He is independent. There are also many voices available, far better than those of Microsoft. But, most are paid and often relatively expensive.

Pré-requis :

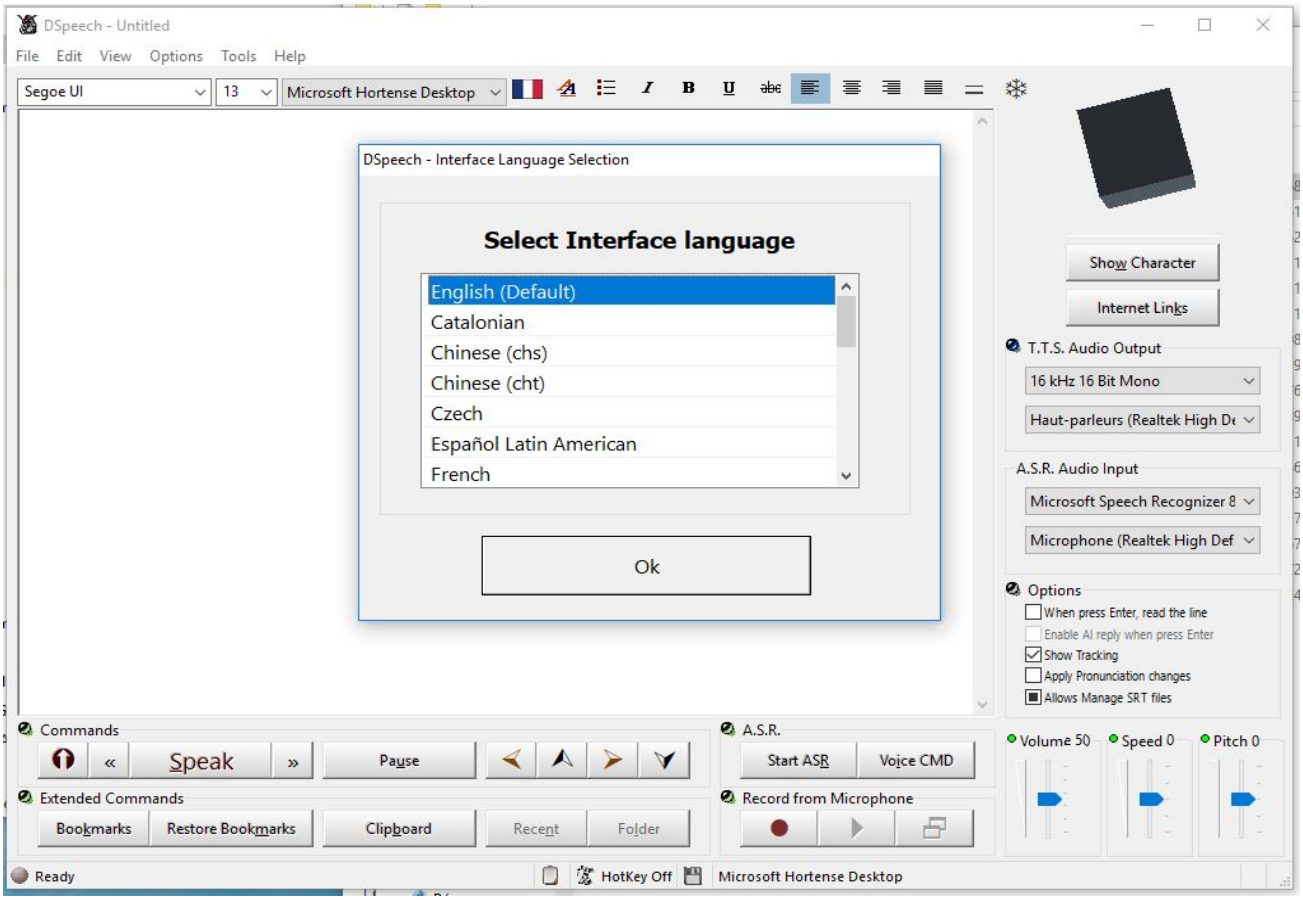
Installation :

After downloading the program on Dimio's website:

1 - Install the program

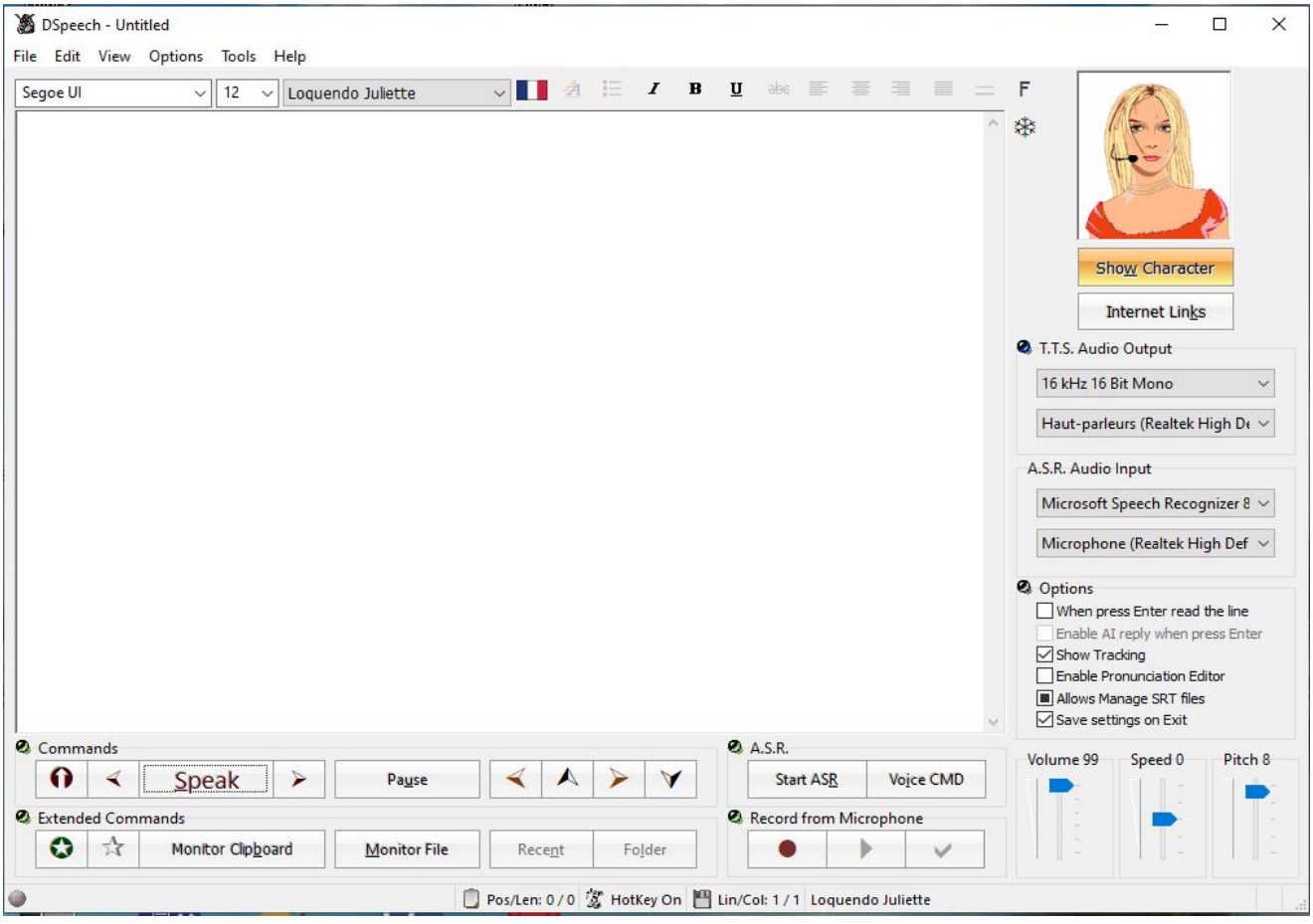
2 - Start the program Dspeech which displays the screen below

The welcome screen is displayed:



Click OK.

The program displays the main screen.



DETAILS:
~~~~~~~~~~

Creation: Michel SAVARD, from the Manual (ENG).txt file

DSpeech (by Dimio)
~~~~~~~~~~~~~~~~~~

The updated version is found to the following Link:

Home: "http://dimio.altervista.org/"

DESCRIPTION:
~~~~~~~~~~~~

DSpeech is a program of TTS (Text To Speech) with functionality of ASR (Automatic Speech Recognition) integrated. And' that is in degree to read to tall voice the written text and to choose the sentences to pronounce according to the vocal answers of the consumer. And' projected specifically to give in rapid way and directed the functions of greater practical utility that are required to the programs of this type, maintaining to the meantime to the least one the invasività and the consumption of resources (it is not installed, it doesn't integrate him in the system, it is light, it sets out in an instant and he/she doesn't write anything in the register).

The principal characteristics of DSpeech are:

1. It allows to save the output in the form of a file Wav or Mp3.
2. It allows to quickly select different voices and to combine her among them to create dialogues among more voices.
3. Entire a system of Vocal Recognition that, through a simple language of script, it allows to create interactive dialogues with the consumer.
4. It allows to shape the voices in independent way.
5. Through special TAG, allows to dynamically vary the characteristics of the voices during the reproduction (speed, volume and frequency), to insert breaks, to emphasize terms or to make the spelling.
6. It allows to capture and to automatically reproduce the content of the ClipBoard.
7. It supports all vocal engines compatible with SAPI 4 and 5.

Entire besides a series of secondary characteristics, among which:

1. To the start, it allows to auto-load the last open file with relative position of reading.
2. It supports the command line and you/he/she can be used then, without graphical user interface, for the creation of audio-books.
3. It allows to specify the format of the audio output, this is able useful venir in very particular situations, when there is the necessity to operate with some files wav of defined characteristics.
4. It allows to create some assemblages inserting, through a special KeyWord, of the fileses wav or mp3. This can be useful to introduce, during the reading, of the particular effects as a hit of cough, a laughter, or also of the you detach musical.
5. When the mp3s are saved, it is possible to specify the quality of the same, in way to be privileged the dimensions or the quality of the result.
6. And' now possible to convert some text in mp3 or wav dividing him/it in file from 5, 10 or 15 minutes each.
7. In the file "CustomTAG.TXT", it is possible to insert some personalized TAG that will appear then in the contextual menu now (for instance the expressive tags of Loquendo).
~~~~~~~~~~~~

TEXT TO SPEECH:

Through the contextual menu (Right Click) it is possible to specify with what voice must be pronounces a date sentence, this it makes the creation of dialogues possible among different voices.
It is likewise possible to insert special TAG that allows to modify the characteristics of the voice while you/he/she is speaking (speed, volume, frequency etc).

OPTIONS AUDIO:

And' possible to specify the card audio in which to redirect the output and, above all, the bitrate of the voices that he is using. It always needs to try to use the same bitrate from the voices in use, in contrary case, losses of quality or the effect can you/they can be had (as if the voice spoke to a can).
In general, the settaggio used by the most greater part of the synthetic voices is: "16 Khzes 16 Mono Bit", while the voices of Microsoft use "22 Khzes 16 Mono Bit."
These settaggis are particularly important when a conversion effects him in file Wav or Mp3.

CONVERSION IN FILE WAV OR MP3:

And' possible to use DSpeech for the conversion of the text in a file Wav or Mp3. If the formed mp3 is chosen, the possibility is given to specify the characteristics of the compression, in fact it can be settata so that to privilege the dimensions, the quality audio of the produced file, or so that to get a balanced thing.

EDITING DI THE TEXT:

Besides the functions of standard editing (find, replace, cut etc.) a particular function is integrated, her "Remove Useless Return Carriage." It serves to eliminate all the present useless carriage return in the text that you/they could limit the fluidity of the reading from the artificial voices. In fact, often understands that a text, for reasons for pagination, contains a series of carriage return that you/they would negatively go to impattare to the quality of reading from the TTS.

CREATION OF AUDIOBOOK:

A series of options they are supported for the creation of audiolibri, in particolar way, the possibility to divide the text in more file of tot minute each. And' also possible to use the manual subdivision of the text in files. In this case, it is necessary to insert the KeyWord #BREAK every time that is wanted to change file.
Through the voice of menu "Append" it is possible to unite more fileses of text one behind the other.
Between the one and the other one the KeyWord will be inserted #automatically BREAK so that to allow the subdivision of the text separate files.
And' also possible to insert a break to the beginning of the text, so that to maintain the compatibility with the readers CD or older mp3.

VOCAL RECOGNITION:
~~~~~~~~~~~~~~~~~~

DSpeech supports a system of vocal recognition that, united to a simple system of script, it makes him/it in degree to create interactive dialogues with the consumer of the type:
CONSUMER: "Computer"
PC: "Ready computer, who are? "
CONSUMER: "Dimio"
PC: "Welcome then"
Etc.

The system of script is very similar to the BASIC, for now the following KeyWordses are supported:

#VOICE Voicename
#I GIVE
#EXIT DO
#LOOP
#RECOGNIZE word1, [word2], [OTHER_WORDS]...
#RECOGNIZE_WITH_TIMEOUT Seconds, word1, [word2], [OTHER_WORDS]...
#IF RECOGNIZED word1, [word2], [OTHER_WORDS]...
#IF TIMEOUT
#END IF
#CALL NomeSub
#Sub NomeSub
#END SUB
#RANDOM
#HOUSES
#END RANDOM
#EXECUTE PathFileOProgram
#OPEN FileToSpeech.txt
#STOP
#BREAK
#PLAY FileName.wav
#WAIT nSeconds
#CLOSE

In the contextual menu (right-click) it is possible to find all these KEYWORDSs with relative Examples.
I am not to explain the syntax considering that you/he/she can deduce her/it in more intuitive way from the examples themselves.
In every case, an example of script for the vocal recognition could be the following:

#VOICE Marco
I am Angelus the computer of Dimitri. You who are?
#DO
 #RECOGNIZE Dimitri, Gloria, OTHER_WORDS
 #IF RECOGNIZED Dimitri
  Angelus waiting for instructions.
  #EXIT DO
 #END IF
 #IF RECOGNIZED Gloria
  You have mistaken computer, yours is that of side.
  #EXIT DO
 #END IF
~~~~~~~~~~~~~~~~~~

```
 #IF RECOGNIZED OTHER_WORDS
  #RANDOM
  #CASE
   Can you repeat please? I have not understood your name.
  #CASE
   What have you said? Can you Repeat?
  #CASE
   I have not understood what you have said, perhaps, simply, your name I don't know him/it.
  #END RANDOM
 #END IF
#LOOP
```

The system of recognition, founds unfortunately entirely for now him on the phonetics English, for which, to make to recognize some words, can be necessary to suit her for the pronunciation English. For instance, to make to recognize to the computer the word "Russia" it needs to write "Rassya."

SHORTCUTS:
~~~~~~~~~~

All the functions of the interface solo in partnership to of the shortcuts of keyboard, the followings special keys are supported besides:

F1 = you Go to the box of editing
F4 = Pause/Resume
F5 = Speak/Stop
F6 or ALT + UP = Speak Previous Line
F7 or ALT + LEFT = Speak Current Line
F8 or ALT + DOWN = Speak Next Line
F9 or ALT + RIGHT = Speak From Cursor
F11 = it Passes to the preceding voice
F12 = it Passes to the following voice
ESC = Stop
ALT+1 = it Increases the volume
ALT+2 = it Decreases the volume
ALT+3 = it Increases the speed
ALT+4 = it Decreases the speed
ALT+5 = it Increases the pitch
ALT+6 = it Decreases the pitch
~~~~~~~~~~

COMMAND LINE:
~~~~~~~~~~~~~

It's possible to specify a file name to open and automatically reproduce. This allows to perform in automatic way a script.

SYNTAX:

DSPEECH.ExE [/Play] [/Speak] [/Wav] [/Mp3] [/Ogg] [/Hidden|/HiddenFix] [FileToSpeech.txt]

COMMAND LINE SAMPLES:

- To open a file:

DSpeech.exe source.txt

- To start a file reproduction:

DSpeech.exe /Play source.txt

- To speak aloud a short sentence:

DSpeech.exe /Speak Hello!

- To convert a text file to mp3:

DSpeech.exe /mp3 source.txt [destination.mp3]

- To convert a text file to ogg:

DSpeech.exe /ogg source.txt [destination.ogg]

- To convert a text file to wav:

DSpeech.exe /wav source.txt [destination.wav]

SYSTEM CONFIGURATION:
~~~~~~~~~~~~~~~~~~~~~

DSpeech asks for a resolution than at least 1024x768.

THE VOICES:
~~~~~~~~~

DSpeech, uses the installed voices in the system, of default, on Windows XP there is only Microsoft SAM (in English besides), while, if MS-SAPI5.1 is installed on Windows NT/2000, disposition they will be had to other two voices (Mike and Mary) also them in English.
The consumers of XP can unload here her from:

http://download.microsoft.com/download/speechSDK/SDK/5.1/WXP/EN-US/Sp5TTIntXP.exe

These last, is surely best of SAM, but their quality is really scarce if compared with voices of third parts (difference is abysmal), for which he recommends to unload of it and to install of it of the others. Unfortunately The best are to payment and, it is not even at times easy to give her for him in legal way. In every case, to the top of the category we find the voices of the manufacturing segentis:
~~~~~~~~~

Acapela (clear and intelligible voices but not the maximum in terms of naturalness).
Cepstral (they are those with the good relationship prezzo/prestazioni, the quality it is not to the same levels of the most expensive voices, but they are valid however).
Loquendo (Very good, especially in terms of naturalness and expressiveness, they also cost so much).
RealSpeak (Surely good).
VoiceWare (Also these are not quite badly, but there are not Italian).
Ivona (Probably the best).

Notes:
~~~~~~

When a file mp3 is inserted in the text, kind if of big dimensions, a small break can be warned between the reproduction of the preceding line and the reproduction of the audio file, this is normal and it doesn't constitute a bug, in every case, when he goes to save the result in the form of file wav or mp3, the break it disappears completely.

For the compression in mp3 the codec is used Blades (www.mp3dev.org), it corresponds to the file "Lame.exe" included in the packet.

CODERS: ~~~~~~~

Dimitrios Coutsoumbas (Dimio)
SKYPE : katafratto
ICQ : 145633952
E-MAIL : cyberdimio@gmail.com
HOME : http://dimio.altervista.org/

BETA-TESTERS:
~~~~~~~~~~~~~

Talksina (talksina@gmail.com)

IMPORTANT TO KNOW:

The program is evolving regularly. Remember to check the exitence of a new version.

ABBREVIATIONS:

UI = User Interface

AI = Artificial Intelligence

Read the History(ENG) .txt file. It contains information retracing the evolution of the program and some descriptions of features.

The features of the main screen.

Let's start with the banner.

Font Size Voice Text characters Alignment of the text 1 2

1 allows to define the thickness of the line spacing

2 displays the character map of the selected font, or displays a custom character table.
see paragraph 7.

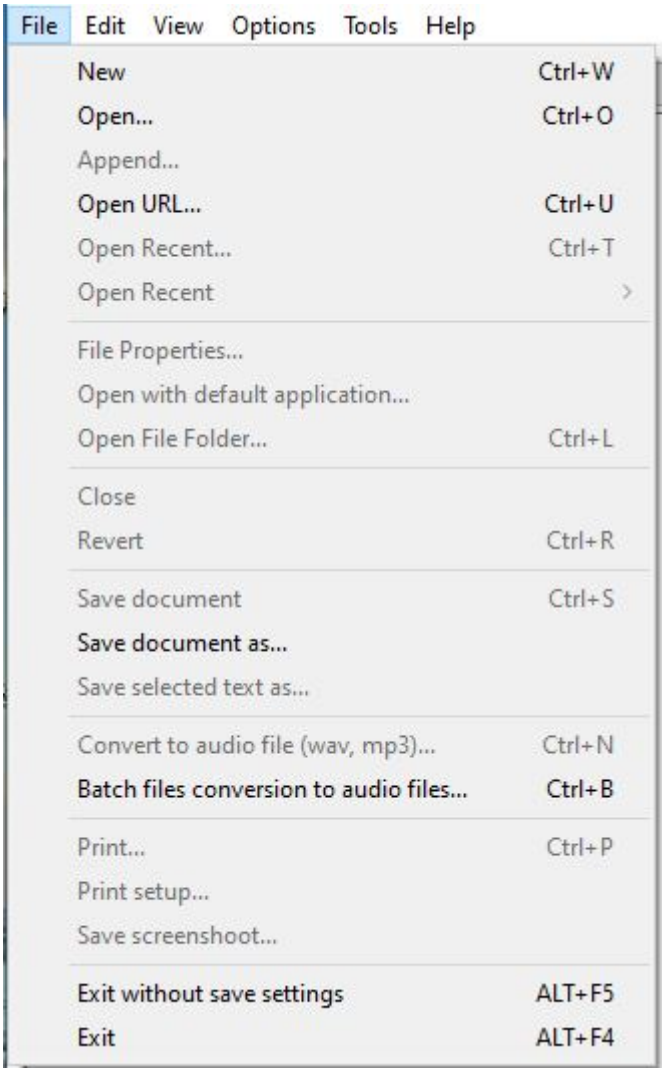
I advise to test all the features.

I use the possibility of recovering the text of Internet pages.

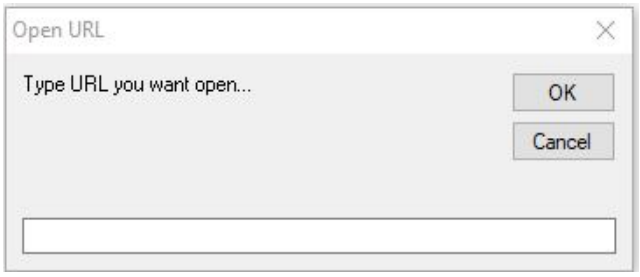
It is advisable to read the texts with a loud enough volume. It is possible to day on the reading speed and the intonation. Do the tests to find the best setting for optimal rendering.

The program is very user-friendly with lots of Info-Bubbles, which will help you in understanding.

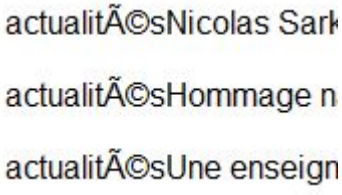
1. The File Menu.



1.1 "Open Url ..." This option allows you to import WEB pages.



The program imports the text of the web page, eliminating the images, and by deleting some of the HTML code. However, there are still characters parasites that should be treated by applying a conversion table.



picture 1

Character processing is achieved by the use of a conversion table, accessible via the "Tools Menu", with the option
"Apply Conversion Table ..."

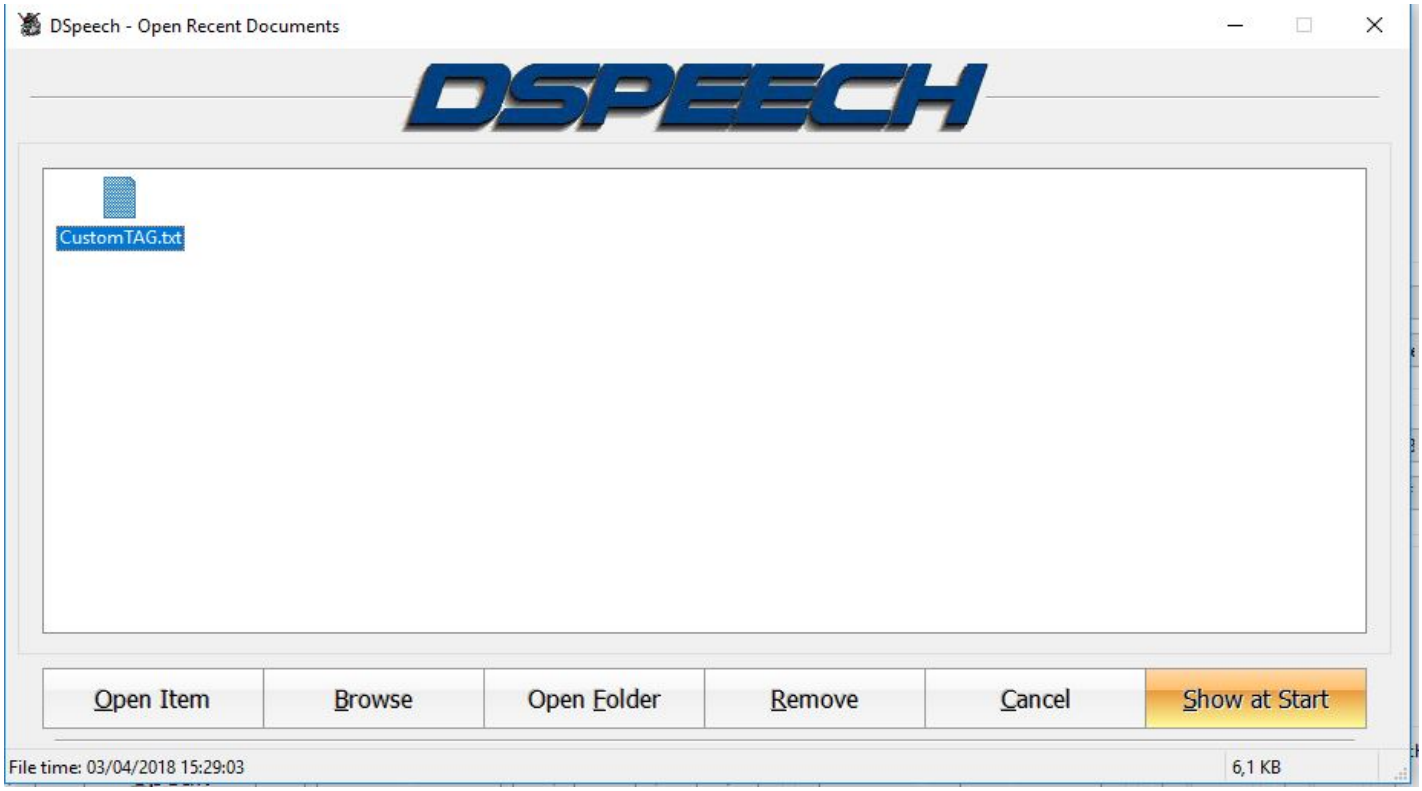
Example of a conversion table "Html.tab", which you can modify according to your needs.

Ã©=é The characters, in red, will be replaced by the letter following the sign =. See picture 1

ì=ë
à€=à
à»=û
î=ë
à´=ô
Â°=°

1.2 The "Open Recent ..." option

The program opens a window in which are listed the last processed files. Different operations are available.

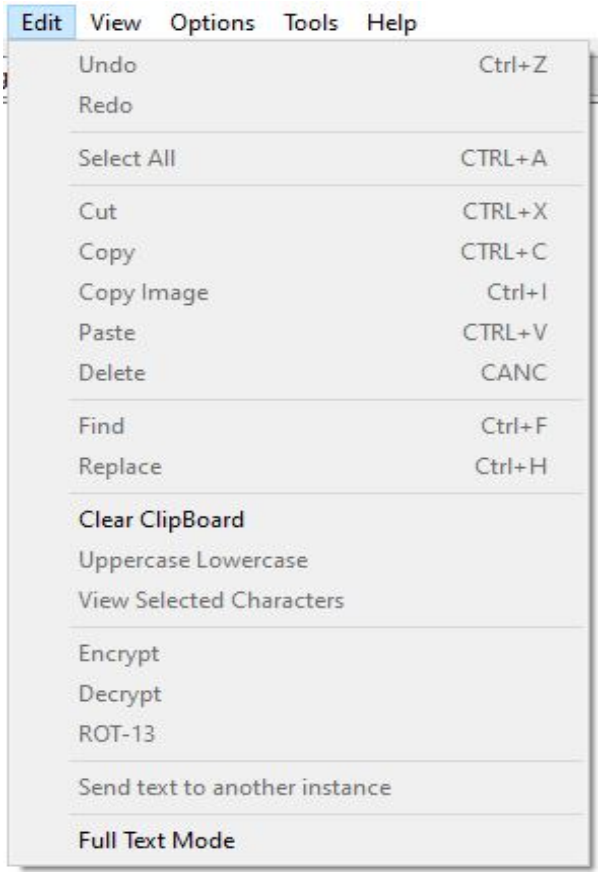


Feel free to test the possibilities.

1.3 Other Options

Familiarize yourself with all the possibilities.

2. The Edit Menu



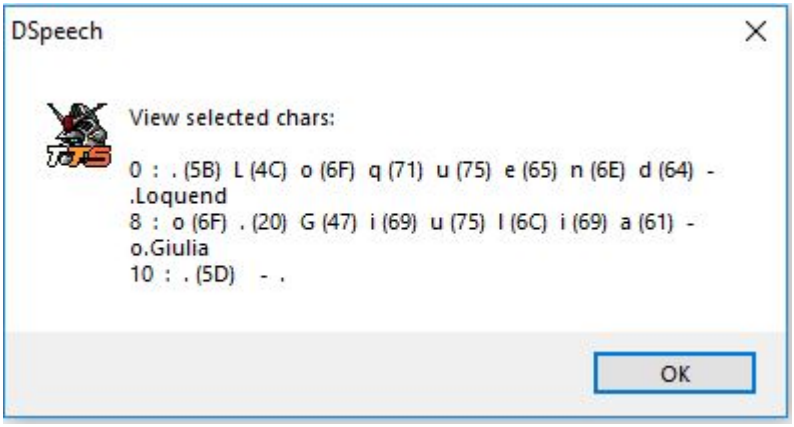
2.1 The function "Copy Bitmap"

The program takes an image of the text contained in the window. Image that you can save as Screenshotnnn.jpg using the "Save Screenshoot" function of the File Menu.

2.2 The function "Uppercase Lowercase"

The selected text will change to upper or lower case.

2.3 The function "View Selected Characters"



The selected text is cut out, character by character, and translated in ASCII code

What is ASCII code ?

The memory of the computer keeps all the data in digital form. There is no way to directly store characters. Each character therefore has its equivalent in numerical code: it is the ASCII code (American Standard Code for Information Interchange - translate "American Standard Code for the Exchange of Information"). The base ASCII code represented the 7-bit characters (ie 128 possible characters, from 0 to 127).

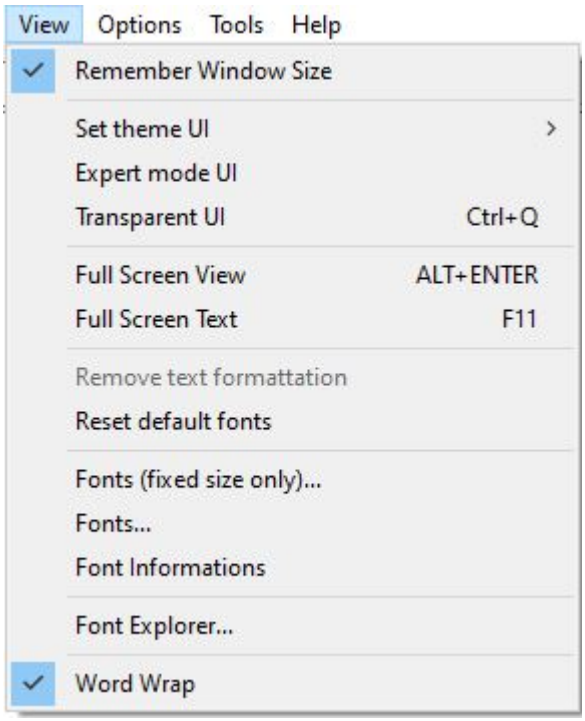
Codes 0 to 31 are not characters. They are called control characters because they allow to do actions such as:
return to line (CR)
Beep sound (BEL)
Codes 65 to 90 represent capital letters
Codes 97 to 122 represent lowercase letters

(Just change the 6th bit to change from upper case to lower case, that is, add 32 to ASCII code in decimal place.)

So:

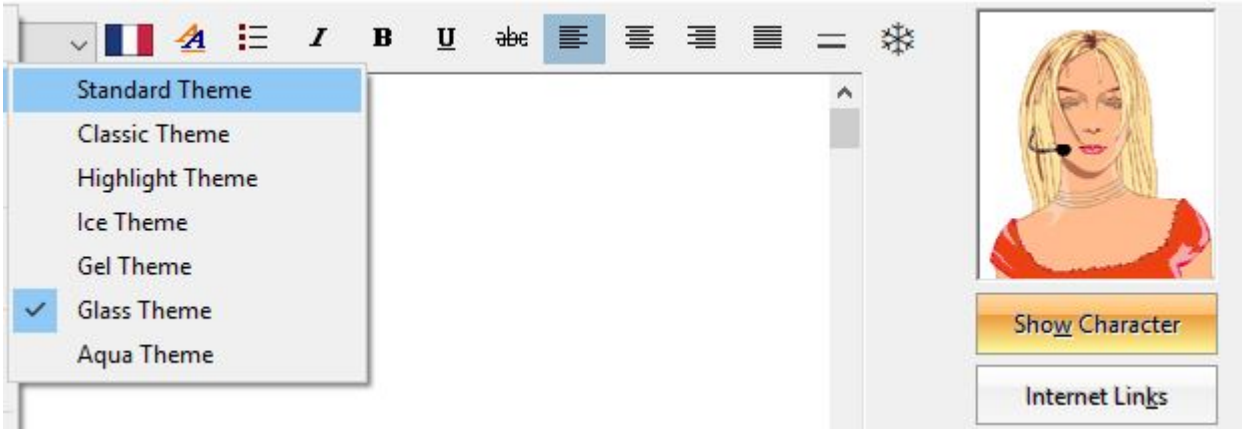
the character "a" whose hexadecimal representation is "61" in lowercase, will have a hexadecimal value "41" in upper case.

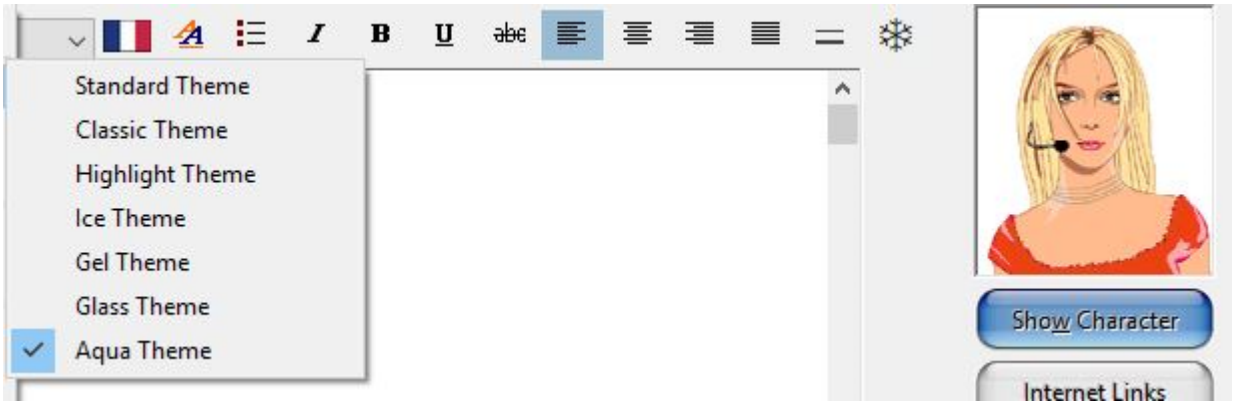
3. The View Menu.



3.1 "Set theme UI" option.

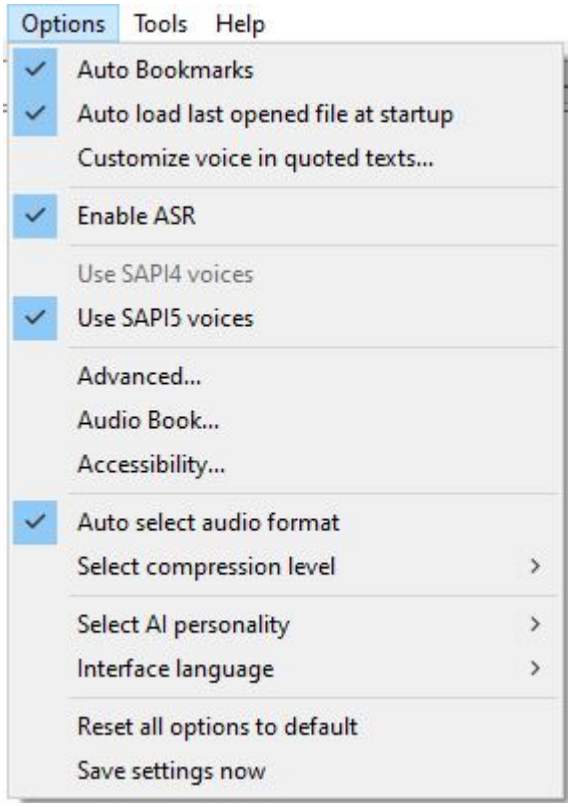
The presentation of the user interface will be adapted according to the choice of theme. It's up to you to find your happiness.





For other options, do the tests and keep what you want.

4. The Options Menu.

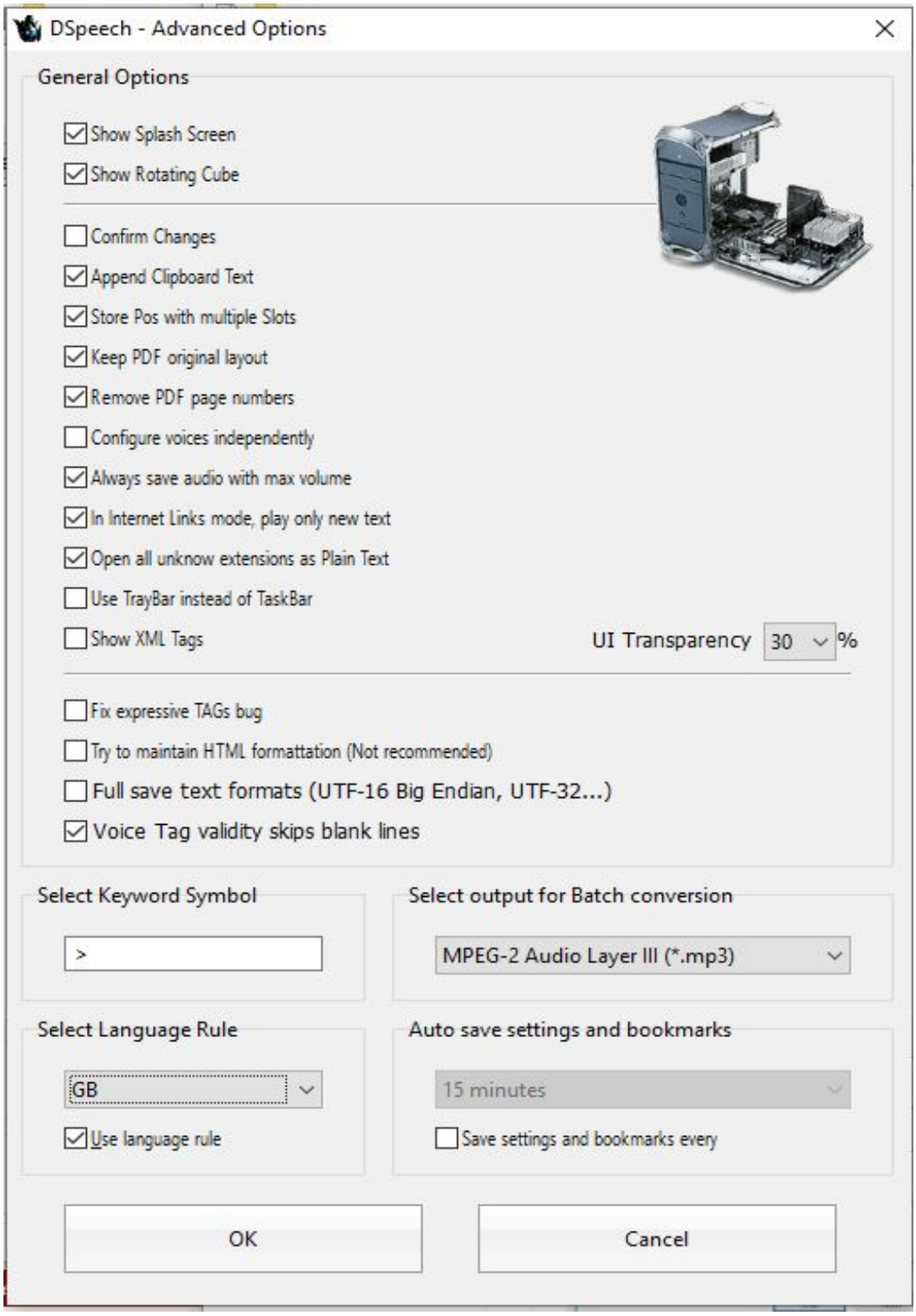


In this section, features are available, which should be tested, to know them well.

For information, SAPI4 voices are no longer usable because of poor quality.

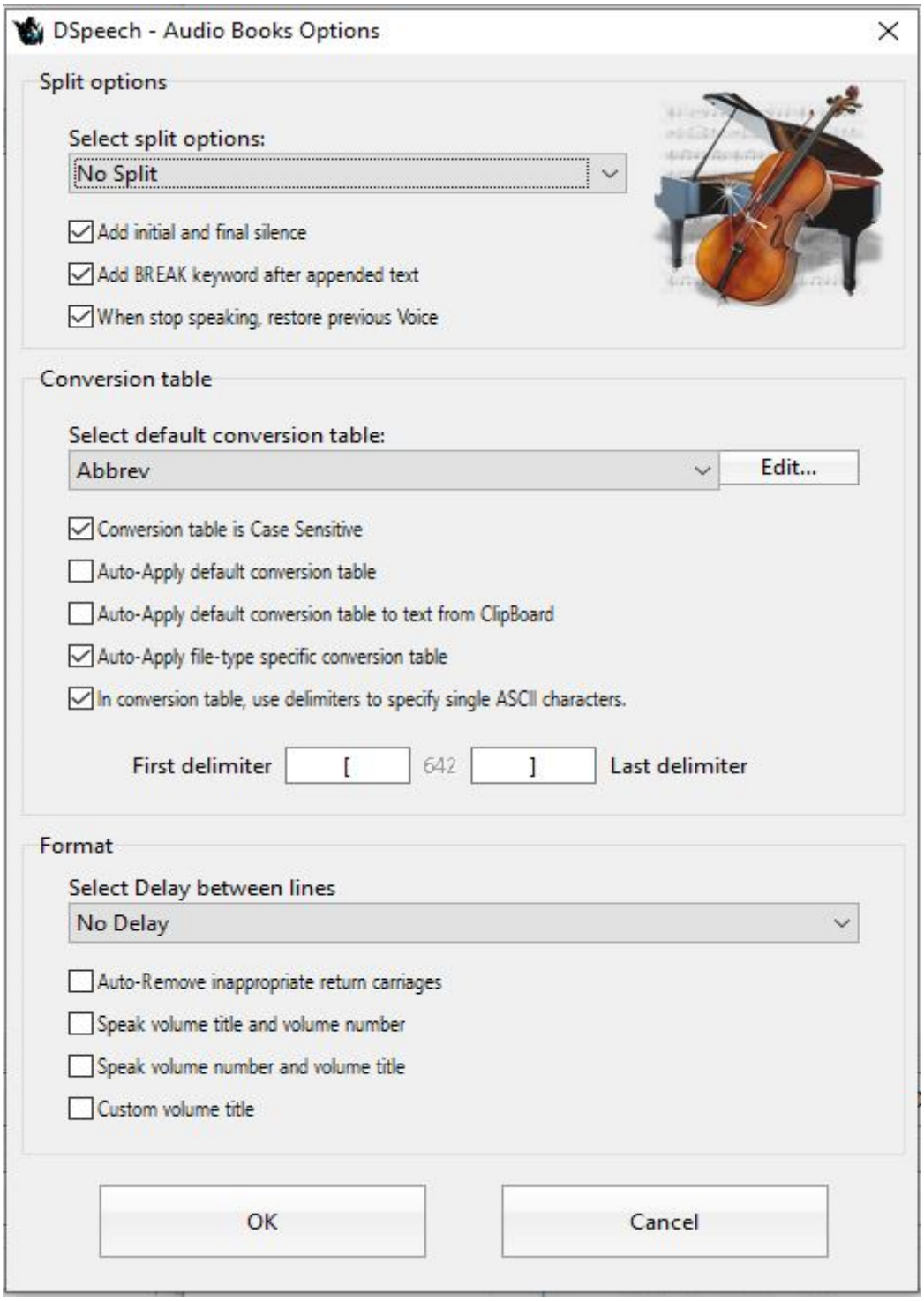
The options "Advanced ...", "Audio Book ..." and "Accessibility ..." are very important. I advise you to test them to exploit at best the program.

4.1 "Advanced Options"



In general, the settings proposed by the program should be retained. However you can always modify these proposals to adapt the program to your needs.

4.2 "Audio Books Options"



4.2.1 Split options.

The facility "when stop speaking restore previous Voice" allows you to keep the settings you set, as long as you use the #VOICE tag in your text, to alternate the voices in converting the text file to audio.

This is related to the option "Voice Tag Validity" of the contextual menu.

See example below, in a dialog.

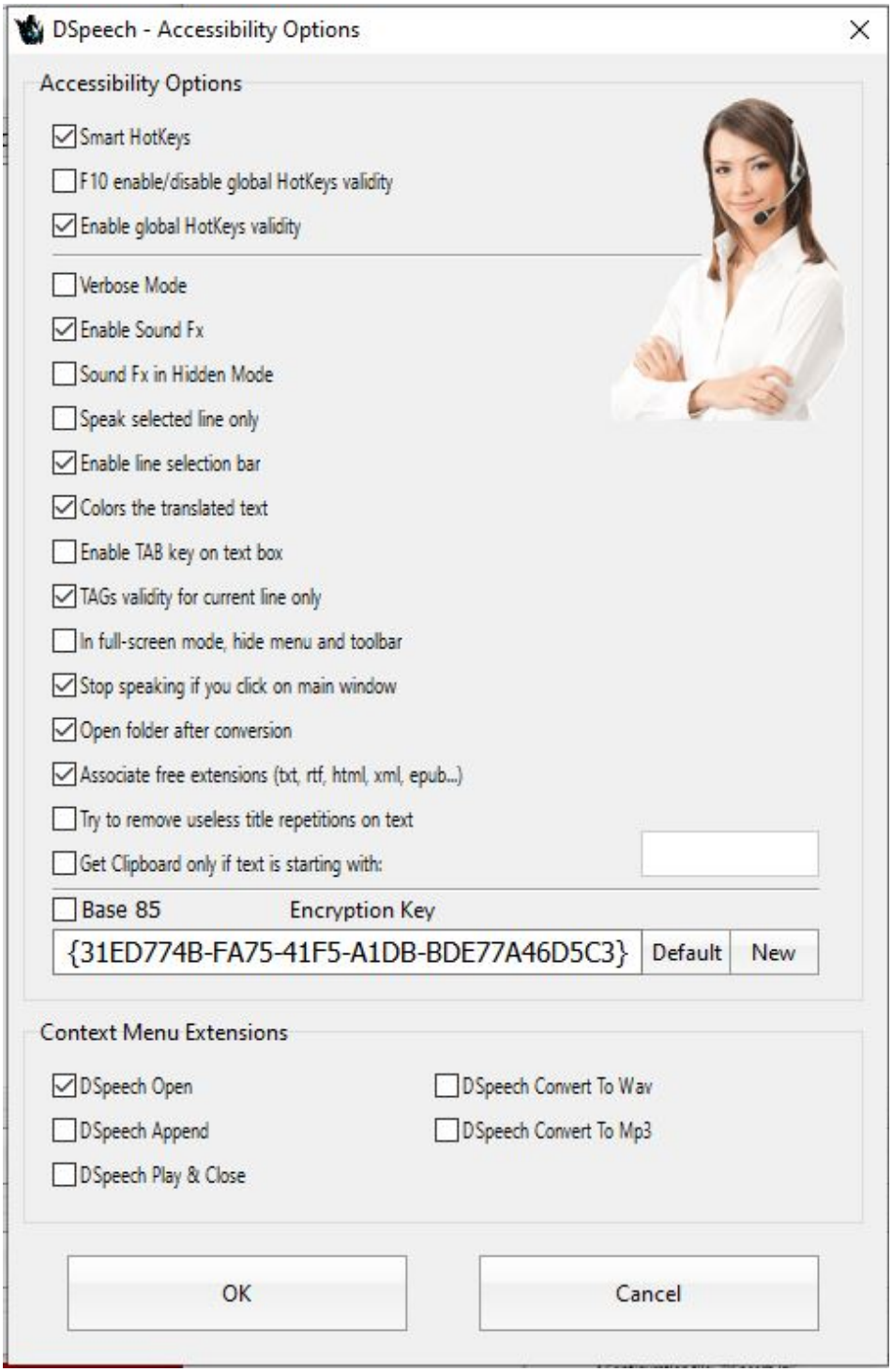
Hello.
I am Julie and i love you.
#VOICE-1 Paul
Hello, i am Paul.
it's a pretty name.
#VOICE-1 Paul
yours is friendly, too, and it looks like you.
thank you.
#VOICE-1 Hortense
be careful! it seems that an idyll is being born.
no worries, mom, I manage.

The reading begins with Julie's voice for 3 lines.
Then the #VOICE-1 Paul tag tells the program that the next line will be read with Paul's voice.
The next line will be read with Julie's voice, until the next tag. And so on. The last line is read with Julie's voice.

4.2.2 "Conversion table is Case Sensitive".

This means that in the case of using a conversion table, the program will respect the character format, including upper and lower case.

4.3 "Accessibility Options".



4.3.1 "Context Menu Extensions".

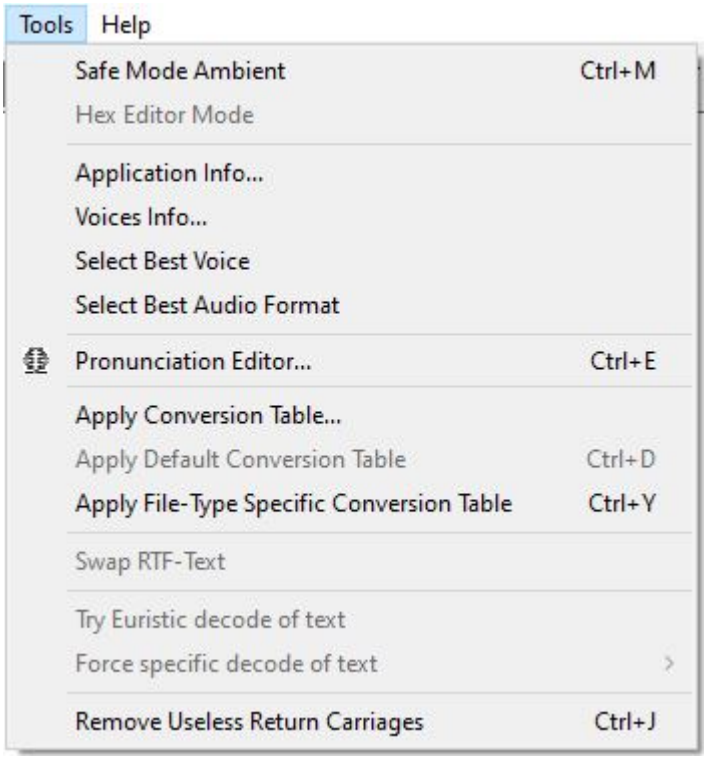
These options can only be changed when the program is running in administrator mode. If this is not the case, you will be informed by a window message pop-up.

From this option, you can directly access certain functions of DSpeech by right-clicking on a file.

the file will be opened directly in DSpeech.

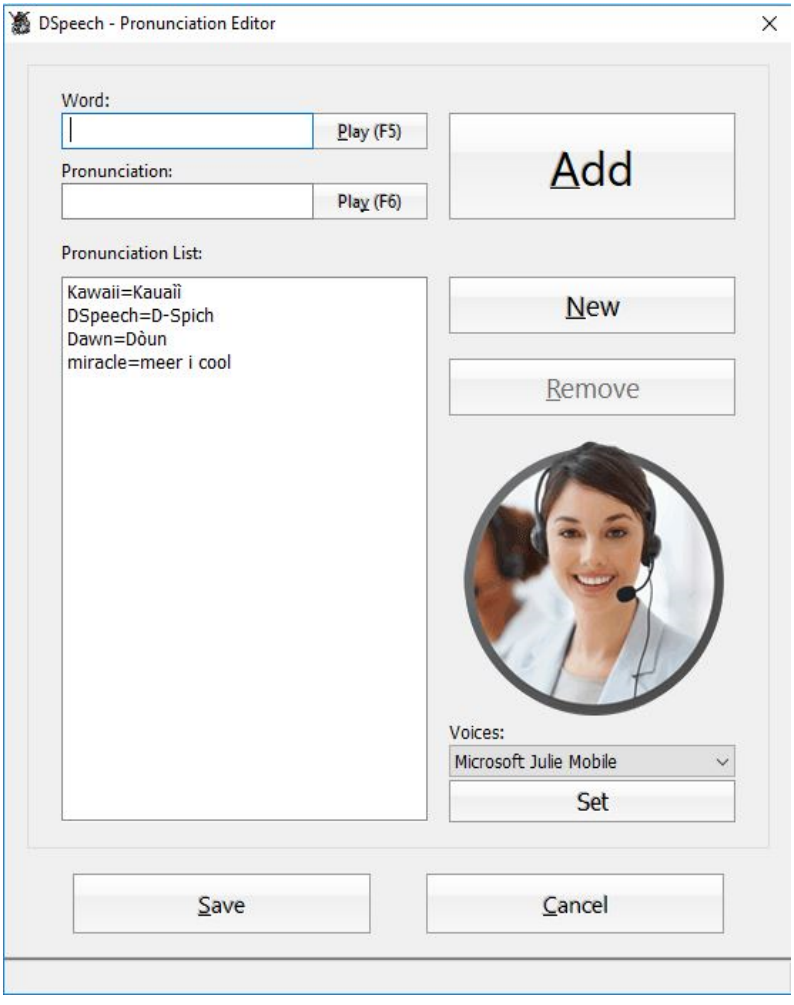
It seems that this obligation is no longer relevant. Because Windows 10 requests permission to run DSpeech, the administrator rights appear to be active.

5. "Tools Menu".



"Hex Editor Mode" function is active only when the program interface is in Expert mode. (see View menu), same for "Force specific decode of text"

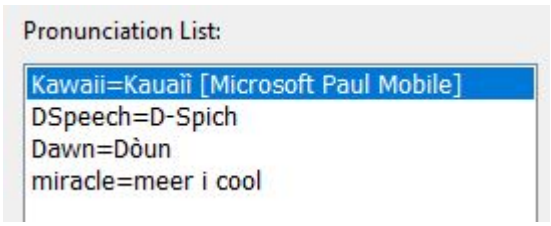
5.1 "Pronunciation Editor...".



This tool allows to improve the pronunciation of words by the program.

It is possible using the function choose to set a particular voice for a given pronunciation.

example:



5.2 "Apply Consersion Table...".

By default the program offers 3 conversion tables:

1- Clear Skype Log.tab deletes the characters between [ ] in the Skype log file

example:
[*]=

2 - Accents.tab allows to correct accented characters and replace with valid characters in the language used.

example:
á=à
â=à
ã=à
ä=à
æ=ae
ê=é
ë=é
í=ì

3 - Epub.tab allows the removal of stray characters when retrieving electronic publications.

example:
 =
 =

for my needs, I created:

html.tab usable during the retrieval of web pages by the program. It converts ISO or HTML characters to frenchfor my needs, I created characters.

txt.tab which converts DOS characters.

odt.tab which removes spurious characters generated by converting ODT files to TXT. This table must be case sensitive.

siecles.tab which translates the abbreviations of century into text readable by the program.

example:
XXIe=twenty first
XXe=twentieth

abbrev.tab which converts certain abbreviations into readable text. To adapt for English speakers.

av.=avant
ap.=après
apr.=après
J.-C.=Jésus-Christ

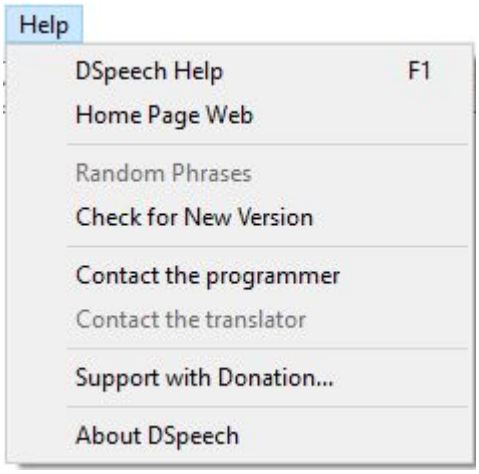
5.3 "Apply Default Conversion Table".

This is the one selected in the Audio Books Options menu

5.4 "Apply File-Type Specific Conversion Table".

Introduction of specific tables for automatic file conversion (for example, the conversion table for ePub files, with this option enabled, automatically applies the "ePUB.tab" table, if present, when opening the file) .

6. "Help" Menu.

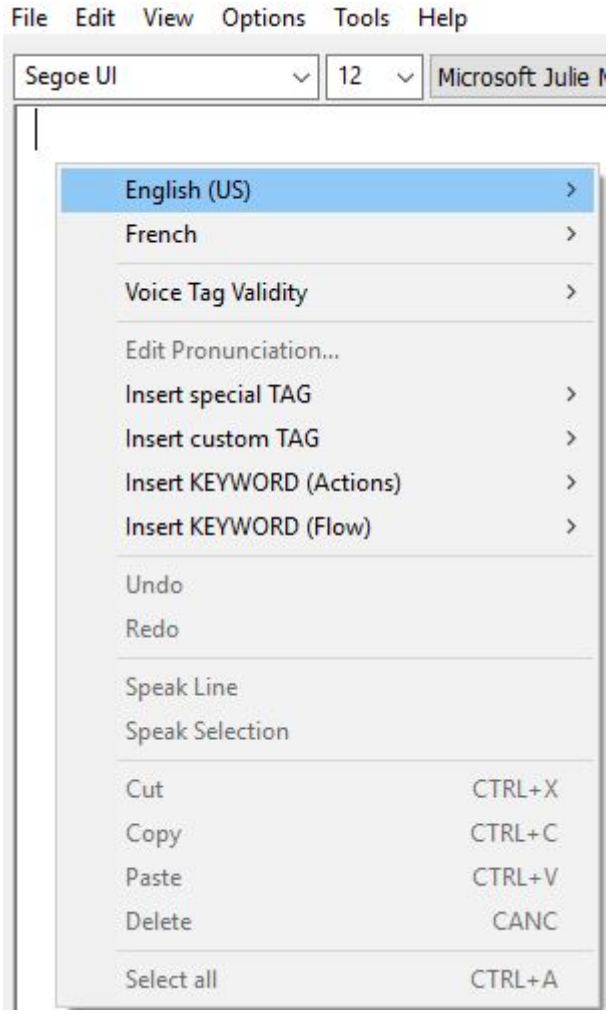


The "DSpeech Help" option gives you information about keyboard shortcuts and other features. To you to discover them.

You can contact the programmer, ie Dimio who is the creator of this tool, for any problem of operation.

7. "Contextual Menu".

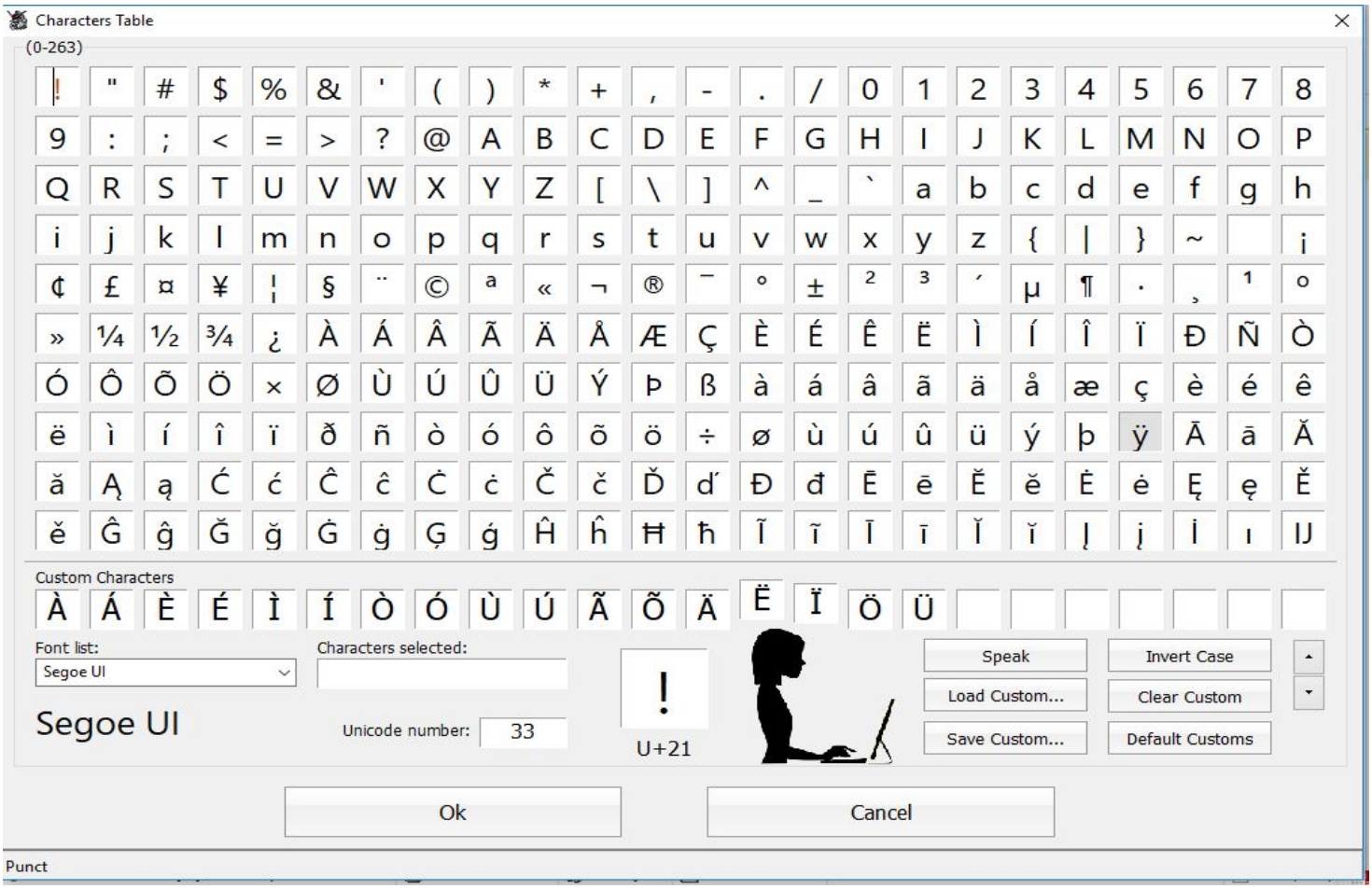
right click on the main window



Discover the possibilities offered by the program.

I like the possibilties that are offered. Between this possibility that we are given to create dialogues that we can add, as commentary in movies in addition to comments made with our own voice.

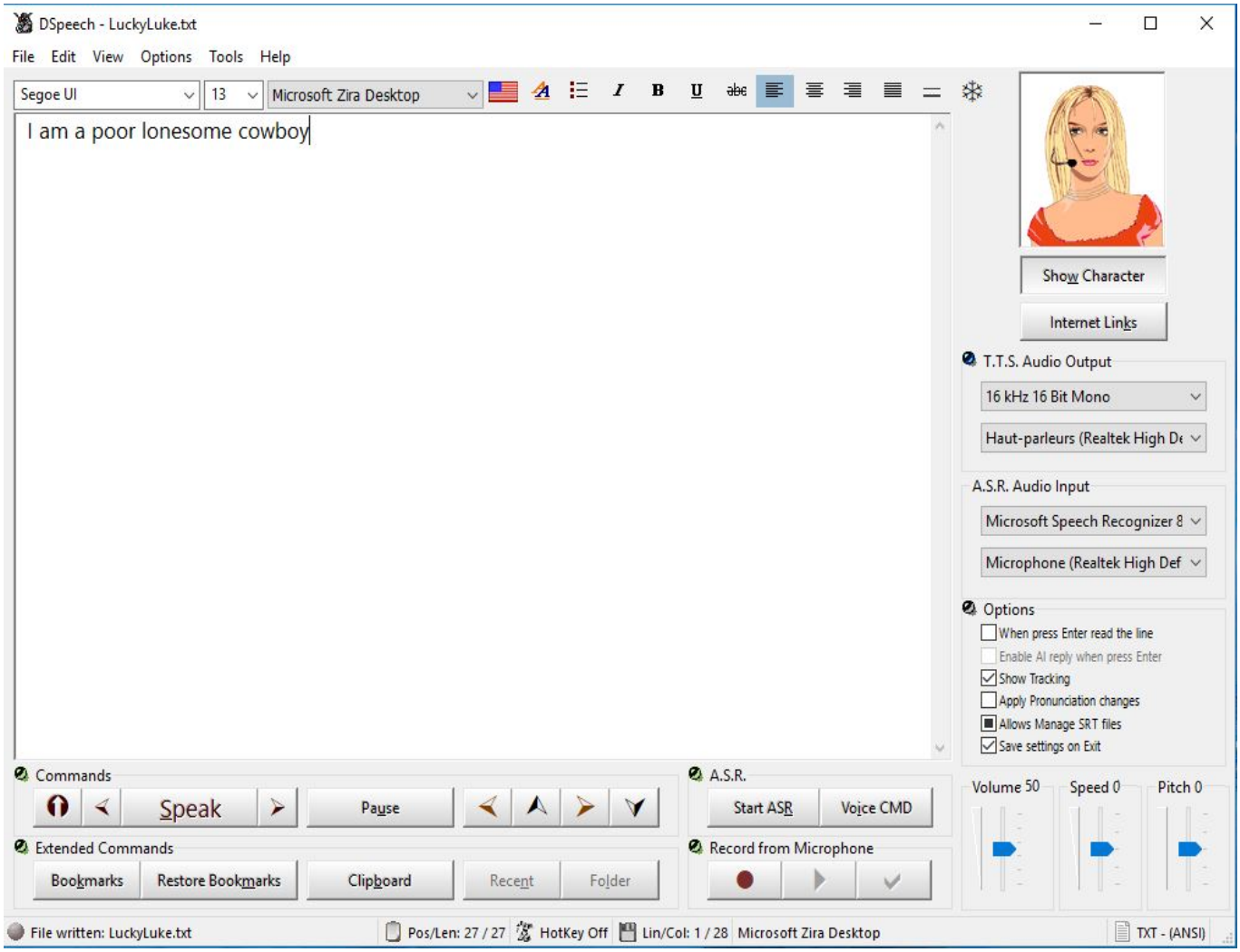
8. The "Characters Table".



In addition to the character map for the font selected on the main screen, you can display a custom table. This is a table in which you can encode the characters
special, like the euro symbol, the character e in the o, or the E accent uppercase. This eliminates the constraint of having to remember combonings of keys, as you can see in the list of characters displayed under the title "Custom character table".

The tables are saved in the DSpeech directory as Charactersxxx.tab.

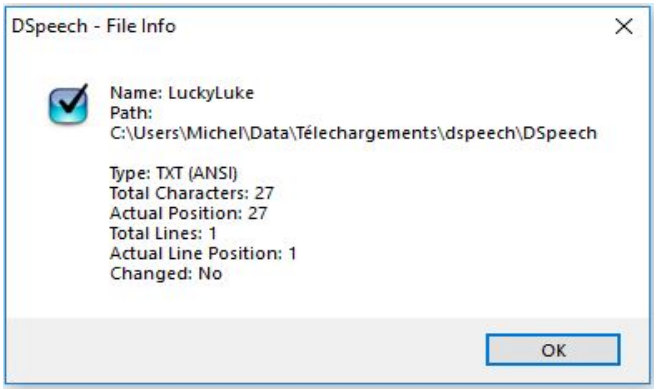
9. More Information



A **B** **C**

Clicking on the icon:

A: You get information about the open text in the main window.



B: You can activate/deactivate HotKeys.

C: You get information about the encoding of the file and its location.