



# **Acapela TTS**

## **for Windows, Mac and Linux\***

### **User's Guide**

**Reference: *AcaMul-user***

\* Acapela TTS for Linux is only available upon request and at specific terms and conditions.

**[www.acapela-group.com](http://www.acapela-group.com), [info@acapela-group.com](mailto:info@acapela-group.com)**

In Belgium: 33 boulevard Dolez, 7000 Mons. Tel +32 (0) 65 77 42 75.

In France: Immeuble Les Erables, rue du Lac, BP 98153, 31681 Labège cedex. Tel + 33 (0) 5 62 24 71 00.

In Sweden: Solna Strandväg 20, PO Box 1328, 171 26 Solna, Tel +46 (0)8 799 86 00

## Document Revisions

Version	Date	Subject	Author(s)
1	27/11/2002	First version BabTTS 1.0	Laurent Hubaut Jean-Michel Reghem
2	19/02/2003	Update	Francois Bataille Laurent Hubaut Jean-Michel Reghem
3	17/03/2003	Update	Laurent Hubaut Jean-Michel Reghem
4	19/05/2003	Update	Laurent Hubaut Jean-Michel Reghem
5	05/07/2004	Update for BrightSpeech 1.3 and Babil 4.1	Laurent Hubaut Jean-Michel Reghem Francois Bataille
6	05/09/2004	Update for BabTTS 1.4	Laurent Hubaut
7	29/08/2005	Update for BabTTS 1.5	François Bataille
8	22/12/2005	Update Acapela Multimedia TTS 6.000 for Windows	François Bataille
9	04/09/2006	Update for Acapela Multimedia 6.000 for Mac OS X	Jean-Michel Reghem
10	19/10/2006	Update	Jean-Michel Reghem
11	02/11/2006	Update	Jean-Michel Reghem
12	19/01/2007	Update for Acapela Multimedia 6.010 for Windows	Laurent Hubaut Jean-Michel Reghem
13	12/02/2007	Minor update (dongle)	Jean-Michel Reghem
14	12/03/2007	Update for Acapela Multimedia 6.010 for Mac OS X	Jean-Michel Reghem
15	26/03/2007	Update Acapela Multimedia 6.010 for Windows	François Bataille
16	11/04/2007	Minor update (page numbering)	Jean-Michel Reghem
17	22/12/2008	Update for Acapela Multimedia 7.000 for Windows: .NET SDK, Signature Bundling, Audio Tag	Laurent Zanoni
18	09/10/2009	Update for Acapela Multimedia 7.5	Sébastien Costanzo
19	01/03/2010	Update for Acapela Multimedia 8.000 for Windows and the page numbering	Daniel Harvengt

20	18/03/2010	Update for Acapela Multimedia 8.000 for Mac OS X	Daniel Harvengt
21	31/05/2010	Update for Acapela Multimedia 8.100 for Mac OS X	Daniel Harvengt
22	16/09/2010	Update for Acapela TTS for Windows 8.200: Added missing <code>\vol\</code> tag description	Daniel Harvengt
23	15/10/2010	Update for Acapela TTS for Windows 8.200	Daniel Harvengt
24	05/01/2011	Update for Acapela TTS for Windows 8.200	Daniel Harvengt
25	18/02/2011	Update for Acapela TTS for Windows 8.201	Daniel Harvengt
26	05/08/2011	Update for Acapela TTS for Windows 8.250	Sébastien Costanzo
27	28/11/2011	Update for Acapela TTS for Windows 9.000	Daniel Harvengt
28	03/02/2012	Update for Acapela TTS for Windows 9.000: Fixed broken <code>\vol\</code> tag description	Daniel Harvengt
29	13/03/2012	Added information about Colibri synthesizer and voices	Daniel Harvengt
30	12/07/2012	Updated information about supported OSes and platforms	Daniel Harvengt
31	19/07/2012	Updated information about supported OSes and platforms	Daniel Harvengt
32	13/08/2012	Lexicon Editor	Sandrine Debray
33	05/02/2013	Update for Acapela TTS 9.020 for Mac	Sandrine Debray
34	18/03/2013	Update for Acapela TTS 9.100 for Windows	Daniel Harvengt
35	22/11/2013	Update for Acapela TTS 9.200 for Windows	Daniel Harvengt
36	19/03/2014	Updated supported operating and supported programming languages info	Daniel Harvengt
37	14/07/2014	Update for Acapela TTS for Windows 9.220	Daniel Harvengt
38	15/10/2014	Update for Acapela TTS 9.220: Added <code>\rmu\</code> tag and <code>\prx\</code> format tag.	Ghilain Florent
39	01/12/2014	Update for Acapela TTS 9.230.	Ghilain Florent
40	26/06/2015	Update for Acapela TTS 9.300.	Ghilain Florent

# Terms and Conditions

By using this software you agree to be bound by the terms and conditions below:

## Software License Agreement

**ACAPELA GROUP SA**

**Bvd Dolez 33 B 7000 MONS BE 0460.125.240 RPM MONS**

**[www.acapela-group.com](http://www.acapela-group.com)**

IMPORTANT - READ BEFORE USING

### Licence Agreement and Limited Warranty

**Please read this license agreement ("license") carefully before pressing the "agree" button. By pressing "agree," you are agreeing to be bound by the terms of this license. If you do not agree to the terms of this license, press "disagree" and (if applicable) return the ACAPELA GROUP software to the dealer where you obtained it. Depending on the terms of delivery of your dealer you are eligible for a (partial) refund.**

#### LICENSE

You may use the programs and manuals on a single computer and copy the programs for back up purposes in support of your use of the programs on a single computer, provided all copyright and other proprietary notices are included on such copy, and true and accurate records of such copy are kept.

**The term "Use", used above, shall not include the right to replicate, redistribute, broadcast or perform in public the sound files generated by the ACAPELA Text-to-speech technologies for commercial purposes. This specific type of use of ACAPELA TTS technology is subject to a separate licensing agreement.**

You may not decompile, disassemble or transfer the programs or manuals in whole or in part, except upon ACAPELA GROUP's prior written approval, or your license and any express warranties granted herein will automatically terminate. In addition, you may not use or copy the programs or manuals, in whole or in part, except as expressly provided for in this license or upon ACAPELA GROUP's prior written approval, or your license will automatically terminate.

#### TERM

This license is effective until terminated. You may terminate it at any time by returning it to ACAPELA GROUP or destroying the programs and any copy you have made of them, and the manuals. This license will also terminate upon conditions set forth elsewhere in this Agreement or if you fail to comply with any term and condition of this Agreement. Unless otherwise stated in this Agreement, you agree upon such termination to return to ACAPELA GROUP or destroy the programs together with the copy, and the manuals.

#### DISK, PROGRAMS AND MANUALS

The programs and the manuals are provided "as is" without warranty of any kind, either express or implied, including but not limited to the implied warranties of merchantability and fitness for a particular purpose. The entire risk as to the quality and performance of the software is with you. Should the software prove defective, you (and not ACAPELA GROUP or any ACAPELA GROUP's representative or any ACAPELA GROUP's licensors) agree to assume the entire cost of all necessary servicing or repair to the computer you are using it on.

ACAPELA GROUP does not warrant that the functions contained in the program will meet your requirements or that the operation of the programs will be uninterrupted or error free. However, ACAPELA GROUP warrants the disk on which the programs are furnished, to be free from defects in materials and workmanship for a period of ninety (90) days from the date of delivery to you as evidenced by a copy of your receipt. ACAPELA GROUP's sole liability in the event of a defective disk shall be to give you a replacement disk. To obtain warranty service or replacement, you must take the disk, or deliver it prepaid, together with a copy of your receipt, to ACAPELA GROUP.

Except to the extent prohibited by applicable law, any implied warranty of merchantability or fitness for a particular purpose on the disk is limited in duration to the duration of this limited warranty.

#### LIMITATIONS OF REMEDIES

In no event will ACAPELA GROUP, ACAPELA GROUP's representatives or ACAPELA GROUP's licensors be liable to you for any damages in excess of the license fee paid by you including, without limitation, any lost profits, business goodwill or other special, incidental or consequential damages arising out of the use or inability to use the disk, the programs or the manuals, or for any claim by any party, even if ACAPELA GROUP, ACAPELA GROUP's representatives or ACAPELA GROUP's licensors have been advised of the possibility of such claims or damages.

#### GENERAL

Some jurisdictions may not allow the exclusion or limitation of incidental or consequential damages, or allow limitations on how long a limited warranty lasts, so the above limitations or exclusions may not apply to you. You may not sell, rent, sublicense, assign or otherwise transfer this agreement or the programs or manuals in any form except upon ACAPELA GROUP's prior written approval. ACAPELA GROUP or ACAPELA GROUP's licensors retain title to and ownership of the programs and manuals and all trade secrets therein at all times. If any part of this Agreement is held invalid or unenforceable, the other parts will remain valid.

You acknowledge that you have read this agreement, understand it and agree to be bound by it. You further agree that unless otherwise expressly agreed to in writing, it is the complete and exclusive statement of the agreement between us which supersedes any demonstration, advertisement, proposal, prior or contemporaneous agreement, oral or written, and any other communications between us or between us and the ACAPELA GROUP representative relating to the subject matter of this agreement.

Acapela TTS for Windows, Mac OS X and Linux and its documentation are copyright Acapela Group 2015. All rights reserved. Infovox iVox is a trademark of Acapela Group. Macintosh, Power Macintosh, Mac, Mac OS and Mac OS X are trademarks of Apple Computer, Inc., registered in the U.S. and other countries. Microsoft and Windows are trademarks of Microsoft Corporation.

## Table of Contents

1. Introduction.....	1
1.1. Acapela's Text-To-Speech Product Range.....	1
1.2. The Technologies.....	1
1.2.1. The Text Processor.....	1
1.2.2. Speech Synthesizers.....	2
1.3. HD, HQ and CO Voices.....	3
1.4. About Manuals.....	3
1.4.1. Differences Between Operating Systems.....	3
1.4.2. Documentation.....	3
1.5. Package Content.....	4
2. Installation of the Software.....	6
2.1. Hardware and Software Requirements.....	6
2.1.1. Hardware Requirements.....	6
2.1.1.1. IBM PC-Compatible Platforms.....	6
2.1.1.2. Apple Platforms.....	6
2.1.1.2.1. Mac OS X Universal Binaries.....	6
2.1.2. Supported Operating Systems.....	7
2.1.3. Supported Programming Languages.....	7
2.2. Installing the Software .....	8
2.2.1. Windows.....	8
2.2.2. Mac OS X.....	11
2.2.3. Linux-based Systems.....	14
3. Testing the TTS Engine .....	15
3.1. Testing the TTS Engine on Windows.....	15
3.2. Testing the TTS Engine on Mac OS X.....	18
3.2.1. Testing with AcaCocoaDemo.....	18
4. Text Tags.....	19
4.1. Exceptions.....	19
4.2. Text Tags Syntax.....	19
4.3. AcaTTS Supported Tags.....	19
4.3.1. \col=wgt=val1,val2[,val3,[val4...]]\.....	19
4.3.2. \emph\.....	20
4.3.3. \equ=preset\.....	20
4.3.4. \equ=val1;val2;val3;val4\.....	20
4.3.5. \ls=number\.....	20
4.3.6. \mrk=number\.....	20
4.3.7. \skipaudio=number\.....	20
4.3.8. \pau=number\.....	21

4.3.9. \paumode=number\.....	21
4.3.10. \pit=number\.....	21
4.3.11. \prn="phonetic string"\.....	21
4.3.12. \prx="phonetic string"\.....	22
4.3.13. \prx=%1nature%word\.....	22
4.3.14. \rms=number\.....	22
4.3.15. \rmw=number\.....	22
4.3.16. \rmu=number\.....	22
4.3.17. \rpit=number\.....	22
4.3.18. \rspd=number\.....	22
4.3.19. \rst\.....	23
4.3.20. \sel=altN\.....	23
4.3.21. \spd=number\.....	23
4.3.22. \vce=key=value\.....	23
4.3.23. \vct=number\.....	24
4.3.24. \vol=number\.....	24
4.3.25. \wrp=number\.....	24
4.4. Audio Tag.....	24
4.4.1. Overview.....	24
4.4.2. Sound files compliance.....	24
4.4.3. Syntax .....	24
4.4.4. Commands.....	25
4.4.4.1. \audio=mix="filepath"\.....	25
4.4.4.2. \audio=offset=X\.....	25
4.4.4.3. \audio=pause\, \audio=resume\ and \audio=stop\.....	25
4.4.4.4. \audio=play\.....	25
4.4.4.5. \audio=play="filepath"[;argument1[;argument2...]]\.....	25
4.4.4.5.1. continue.....	25
4.4.4.5.2. duration=timeduration.....	26
4.4.4.5.3. until=timeposition.....	26
4.4.4.6. \audio=repeat=status\.....	27
4.4.4.7. \audio=volume=percentage\.....	27
4.4.5. Limitations.....	27
5. Acapela TTS License .....	28
5.1. The License Identity.....	28
5.1.1. Site Authorization.....	28
5.1.2. Dongle Authorization .....	28
5.1.3. Bundling.....	28
5.2. The License File.....	28
5.3. The License Manager.....	29

---

5.3.1. Microsoft Windows.....	29
5.3.1.1. Getting a Free 30 Days Trial License.....	30
5.3.1.2. Getting a Developer License (Computer-based).....	30
5.3.1.3. Other Operations.....	30
5.3.2. Mac OS X.....	31
5.3.2.1. Getting a Free 30 Days Trial License.....	31
5.3.2.2. Getting a Developer License (Computer-based).....	31
5.3.3. Linux-based Systems.....	32
5.3.3.1. Getting a Commercial License.....	32
6. User Lexicons .....	33
6.1. Definition.....	33
6.2. Features.....	34
6.3. Using The Pronunciation Editor (Windows).....	34
6.4. Using The Java-based Pronunciation Editor (Windows and Mac OS X).....	36
6.4.1. Location.....	36
6.4.2. Requirements.....	36
Windows and Linux-based Systems:.....	37
Mac OS X Systems:.....	37
6.4.3. Main Window.....	37
6.4.4. The Add/Modify Window.....	38
6.4.4.1. New Word Frame.....	38
6.4.4.2. New Pronunciation Frame.....	38
6.4.4.3. The Phonetic Table .....	38
6.4.5. Default Lexicons.....	39
6.4.5.1. Adding a Lexicon to The Default List.....	40
6.4.6. Managing Open Lexicons.....	40
6.4.7. Importing/Exporting Lexicons.....	41
6.4.7.1. Importing Abb/Exc Files.....	41
6.4.7.2. Importing Text Files .....	42
7. Voice Manager (Windows).....	44
7.1. Definition.....	44
7.2. The Voice Manager.....	44
7.2.1. Introduction .....	44
7.2.2. The Selection Tree.....	45
7.2.3. The Parameters Tabs .....	45
7.2.3.1. Info Tab.....	45
7.2.3.2. General Tab.....	48
7.2.3.3. Pauses Tab.....	49
7.2.3.4. Reading Mode Tab.....	50
7.2.3.5. Equalizer Tab.....	51

---

---

7.2.3.6. Other Controls.....	52
8. SpeechCreator (Windows).....	53

# 1. Introduction

## 1.1. Acapela's Text-To-Speech Product Range

Acapela Group is a global player on the speech market providing a wide range of multilingual Text-To-Speech (TTS) and Automatic Speech Recognition (ASR) solutions covering most of the market needs.

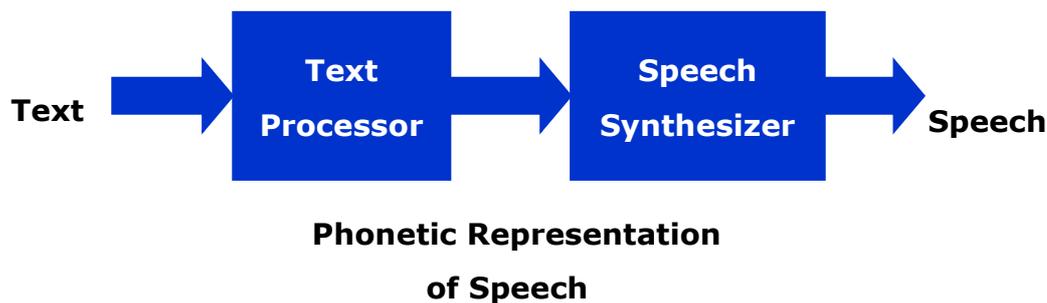
Please visit our website <http://www.acapela-group.com/> for more information on our wide range of products.

## 1.2. The Technologies

Generally speaking, a Text-to-Speech algorithm is composed of two main software modules:

- The text processor (Natural Language Processor — NLP) converts the text into a phonetic representation together with required information to generate the appropriate intonation;
- The speech synthesizer converts this phonetic representation of text into a speech signal.

These two software modules use databases to perform their assigned operation. All databases are language dependent. Databases used by the speech synthesizer are also language and voice dependent.



### 1.2.1. The Text Processor

The Text processor performs the analysis of the input text and converts this text into a sequence of phonemes combined with prosodic information. This software code is language independent for most of the Indo-European languages. The databases contain all information specific to the language.

The Text Processor module is composed of:

- Pre-processor: analyses the text and defines each text unit (word), it extracts the numbers, dates, abbreviations.
- Phonetizer and dictionaries: converts each text unit received from the pre-processor into a sequence of phonemes.
- Prosody generation: maps the rhythm and the intonation onto the phonetic representation of the text.

This block transmits a phonetic string (phonemes, intonation and duration) to the speech synthesizer block.

To ensure a correct pronunciation of special words, access to USER dictionaries is implemented. The User Dictionary brings huge flexibility to the application developer and the final user. It allows the TTS to handle some linguistic info for specific words: proper names, specific professional terminology, abbreviations, acronyms, etc. (see section 4 for more details).

### 1.2.2. Speech Synthesizers

---

During the last decades, the quality of TTS systems has moved from a robotic speech to a warm, natural sounding voice. Our wide range of products is based on three technologies:

- **Diphone concatenation (HD Voices):** provides high quality speech synthesis based on technique like the Acapela Group's patented Multi Band Resynthesis Overlap Add (MBROLA) systems.

This approach allows spectral smoothing of the concatenation points, producing a much more natural voice than with other concatenative systems.

The voices using this technology are marked with letters *HD* (High Density). With this technology, only one instance of each speech unit (diphone) is stored in the database and signal processing (MBROLA) is applied to the units to modify its duration and pitch curve.

- **Unit selection (HQ and alternative voices):** a library of pre-recorded human speech units generates a clear and natural-sounding voice, achieving a superb quality level never reached before.

This new-generation TTS solution significantly improves intelligibility and listening comfort of speech output.

A *HQ* (High Quality) suffix is added to the name of the voices using a synthesizer based on Unit Selection Speech Synthesis.

The idea is to find a unit in the database that almost matches the right duration and pitch curve and to apply to it as little as possible signal processing.

Alternative voices – based on the same technology – with a voice database size smaller than their *HQ* voice counter-part – using a bigger compression rate – are also available on request.

- **HMM synthesis (CO Voices):** Colibri is the name of the brand new Speech Synthesizer of Acapela.

This statistical parametric speech synthesis system is based on Hidden Markov Models (HMMs).

In this synthesizer, context-dependent HMMs are trained from databases of natural speech, allowing to generate speech waveforms from the HMMs themselves based on the maximum likelihood criterion.

This system offers the ability to model different styles without requiring the recording of very large databases.

Even if the overall quality is not at the same level as the Unit Selection voices are, Colibri voices are much smaller (less than 1 MB), have a better intelligibility and the advantage of being consistent in quality. Due to his parametrical aspect, Colibri voices could be manipulated with more flexibility than Unit Selection-based voices, including pitch and diphone duration modification.

---

## 1.3. HD, HQ and CO Voices

---

While **HD**, **HQ** and **CO** voices are based on different technologies, they are based on the same software architecture, called "AcaTTS".

Hence, the development of an application integrating **HD**, **HQ** or **CO** voices is identical from the developer's point of view. Only the selection of the voice will determine the underlying technology.

---

## 1.4. About Manuals

---

### 1.4.1. Differences Between Operating Systems

---

Acapela TTS contains some small differences between Microsoft Windows, Apple Mac OS X and Linux-based versions. In the following sections, such specificities are highlighted by 3 different icons:

-  for the Microsoft Windows family.
-  for the Apple Mac OS X family.
-  for Linux-based systems.

---

### 1.4.2. Documentation

---

The documentation is divided in several documents:

- **User Guide**, this document, describing the installation procedure (section 2) after a brief introduction to the products (section 1). The Software Development Kit comes with a "Type-and-Talk" demonstration (Microsoft Windows only) which is described in section 3. Section 4 explains how licenses are managed by Acapela TTS. Section 5 introduces the tags you can embed in text. Section 6 describes how to handle specific pronunciations through out the Lexicon Editor. Finally, Section 7 describes how to personalize voices to your needs throughout the Voice Manager (Microsoft Windows only).
- **Developer's Guide**
  - The main document is the SDK Developer's Guide ([ACAMUL-SDK](#)) which describes the AcaTTS API. It initiates the developer to the basic mechanism involved in the development of applications in section 2 (Getting Started), section 3 (Creating your first Application with AcaTTS API), section 4 (Using Callback). Section 5 explains how you can personalize the produced speech. Finally, section 6 gives detailed information on each function and structure available in the proprietary API AcaTTS.
  - The License Bundling Guide ([ACAMUL-BUNDLING](#)) and the Custom Installer Guidelines ([ACAMUL-CCD](#) for Windows and [ACAMUL-CCD-OSX](#) for Mac OS X) describe how to bundle your application using Acapela TTS.
  -  SAPI 4 Supported Interfaces ([ACAMUL-SAPI4](#)) and SAPI 5 Supported XML tags ([ACAMUL-SAPI5](#)) giving you some information for using Acapela TTS with Microsoft SAPI 4

and SAPI 5.

-  .NET assembly Reference Guide (**ACA MUL-NET SDK**) describing how to use the AcaTTS .NET assembly wrapper in the .NET framework (VB.NET, C#...)

## 1.5. Package Content

---

The components of the AcaTTS SDK are the following:

- Programming interfaces:
  - **Proprietary AcaTTS API** usable in C/C++ working with the most common compilers (Visual C++, C++Builder, GCC, ...)
  -  .NET framework assembly wrapper for AcaTTS, for VB.NET, C#, Delphi.NET
  -  Microsoft **SAPI 4.0a** standard interface
  -  Microsoft **SAPI 5.1** standard interface
- One or several engines (voices) each composed of:
  - A **Language Module** (i.e. English, French) able to transform a text into a phonetic string.
  - A **Voice Module** that converts this phonetic string into speech. Each voice is defined by the *speaker name* and the *sampling rate* allowing the use of the SDK in a wide range of applications.
- Several development tools:
  -  **AcaTTSDemo**: Type-and-Talk application that allows users to quickly verify the correct installation of the software.
  -  **AcaCocoaDemo**: Type-and-Talk application that allows users to quickly verify the correct installation of the software.
  - **The Lexicon Editor** allow user to create "user dictionaries" in order to change the pronunciation of words according to their needs.
  -  **The Pronunciation Editor** also allows user to create "user dictionaries" in order to change the pronunciation of words according to their needs.

-  **The Voice Manager** allows creating new customized voices, based on original ones but with their own set of characteristics (like rate of speech, pause durations, etc.) according to your specific needs.
- **The License Manager** is used for providing "License Identities" to users for generating purchase requests and managing your installed licenses.
- **The Signature Bundling tool** to create licenses for deployment (requires a specific contract).

## 2. Installation of the Software

The SDK is generally distributed on DVD containing the required components of the TTS. In the most common case, you will receive either a HQ package (several DVD-DL), a HM package, a LF package or a Colibri package. Most packages also ship HD voices.

### 2.1. Hardware and Software Requirements

#### 2.1.1. Hardware Requirements

##### 2.1.1.1. IBM PC-Compatible Platforms

	<b>HD Tehnology Requirements</b>	<b>HQ Tehnology Requirements</b>	<b>CO Tehnology Requirements</b>
Processor	Pentium III (800 Mhz)	Pentium III (800 Mhz)	Pentium III (800 Mhz)
Memory	16 MB	At least 128MB (it depends on the application)	16 MB
Free Disk Space (per voice)	30 MB	± 100 → 150 MB (could be bigger for some languages)	6-12 MB

A 16-bit digital audio board or a 16-bit external audio adapter supported by the operating system must be installed for multimedia applications. Minimum memory has to be

##### 2.1.1.2. Apple Platforms

	<b>HD Tehnology Requirements</b>	<b>HQ Tehnology Requirements</b>	<b>CO Tehnology Requirements</b>
Processor	PPC: Not supported anymore starting with Acapela TTS 8.000 for Mac and higher  i386: All Apple MacIntel Platforms	PPC: Not supported anymore starting with Acapela TTS 8.000 for Mac and higher  i386: All Apple MacIntel Platforms	PPC: Not supported anymore starting with Acapela TTS 8.000 for Mac and higher  i386: All Apple MacIntel Platforms
Memory	512 MB	512 MB	512 MB
Free Disk Space (per voice)	30 MB	± 100 → 150 MB (could be bigger for some languages)	6-12 MB

##### 2.1.1.2.1. Mac OS X Universal Binaries

The executable files and modules in the Mac OS X version of Acapela TTS are "Universal Binaries" which contain 2 processor architectures (i386 and x86\_64) and can therefore be used natively on both 32- and 64-bit Mac Intel platforms (MacBook, MacBook Pro, etc.).

## 2.1.2. Supported Operating Systems

---

-  Microsoft Windows:
  - XP SP3 (x86 edition only)
  - Vista RTM, SP1 and SP2 (x86 and x64 editions)
  - 7 RTM and SP1 (x86 and x64 editions)
  - 8 RTM (x86 and x64 editions)
  - 8.1 RTM (x86 and x64 editions)
-  Microsoft Windows Server:
  - 2003 RTM and R2 (x86 and x64 versions)
  - 2008 RTM (x86 and x64 editions)
  - 2008 R2 (x64 edition)
-  Apple i386 and x86-64 platforms (Mac Intel):
  - Mac OS X 10.6.8 and higher

## 2.1.3. Supported Programming Languages

---

-   C/C++ interface (Microsoft Visual C++ 2008, GCC 3.2 and higher) and all languages using standard WINAPI Dynamic Link Libraries (DLL).
-  Microsoft SAPI 5.1.
-  VB .NET and C# using the SpeechLib component of Microsoft SAPI5.1.
-  VB .NET and C# using the AcaTTS .NET assembly wrapper.
-  C/C++/Objective-C/Objective-C++ interface with XCode 3.2 and higher, GCC 4.0 and higher.

-  Carbon speech synthesis API (also called Speech Synthesis Manager) in the ApplicationServices Framework (C/C++/Objective-C/Objective-C++).
-  NSSpeechSynthesize object of the Cocoa Speech Synthesis API in the AppKit Framework (Objective-C and Java).

## 2.2. Installing the Software

---

The SDK comes with a fully automatic installation procedure on Windows (custom "AcaWizard" installer) and on Mac OS X (through the Apple Installer). It is highly recommended to use this software for installing the TTS.

If your application needs specific installation procedures (i.e. for packaging), please refer to the custom installer guidelines ([AcaMul-CCD](#) for Windows or [AcaMul-CCD-OSX](#) for Mac OS X).

### 2.2.1. Windows

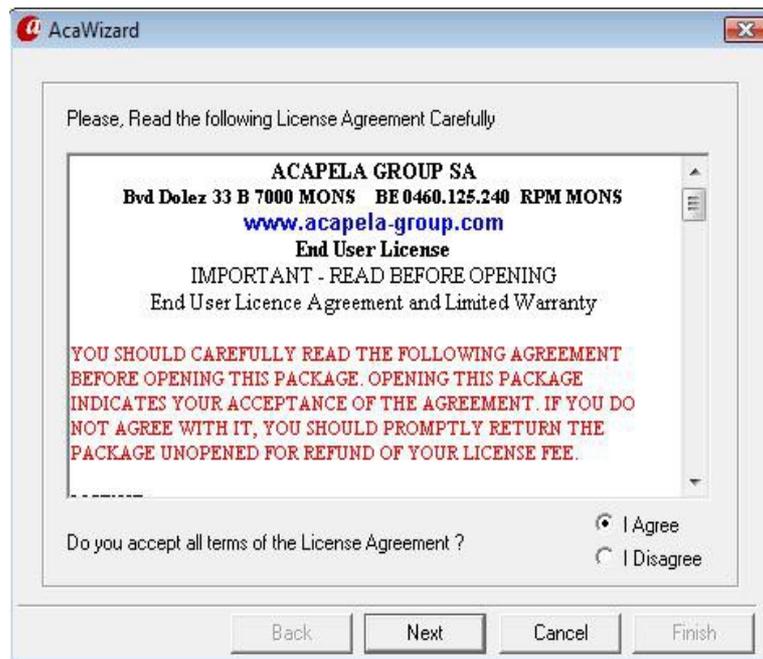
---

When you insert the CD, the AcaWizard program automatically starts.

You can also start it by yourself by double-clicking on *AcaWizard.exe*.



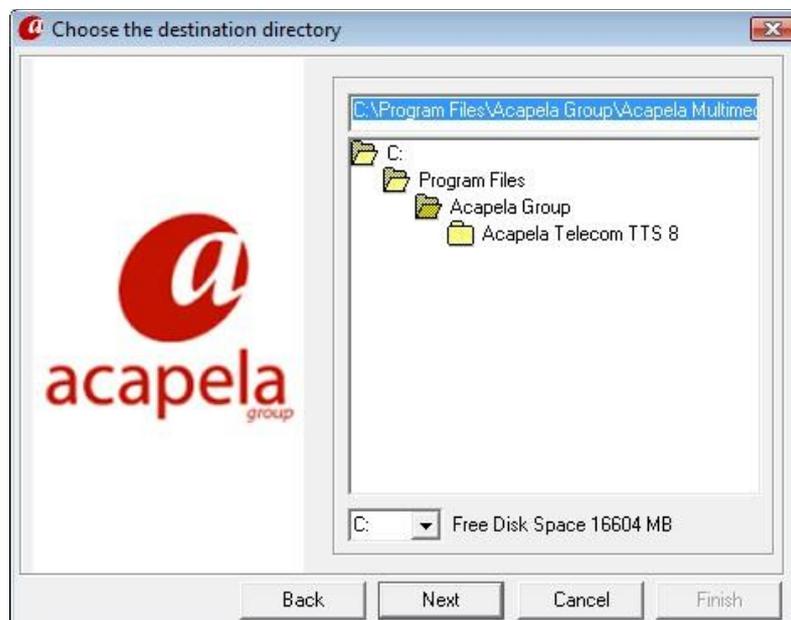
Click Next to go to the next screen.



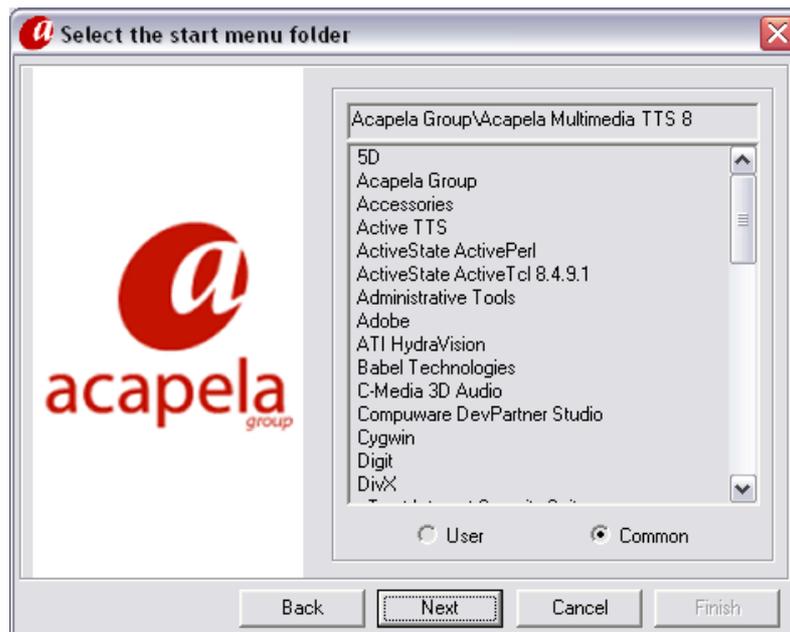
Carefully read the license agreement, if you agree, select "I agree" and click Next.

In the next two screens, you can choose

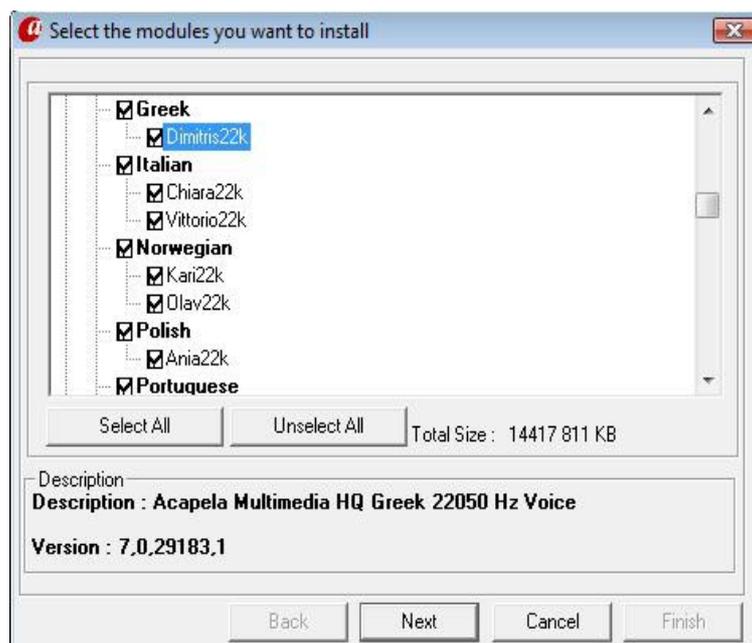
- The location of the directory where you want to install the software. Note that all further installation of modules will use the same directory structure.



- The start menu folder where you want to insert link to the different tools.

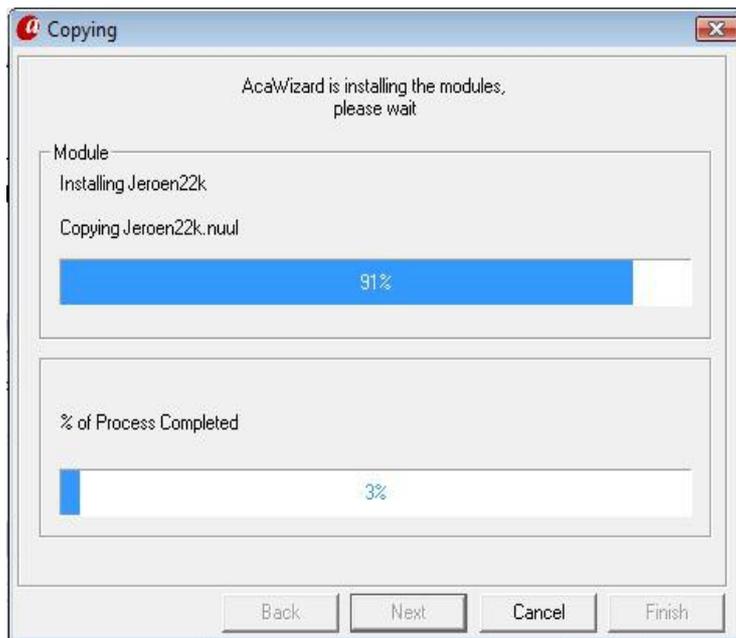


You can then select the modules you want to install (by selecting the corresponding check box). By default, all components are selected.

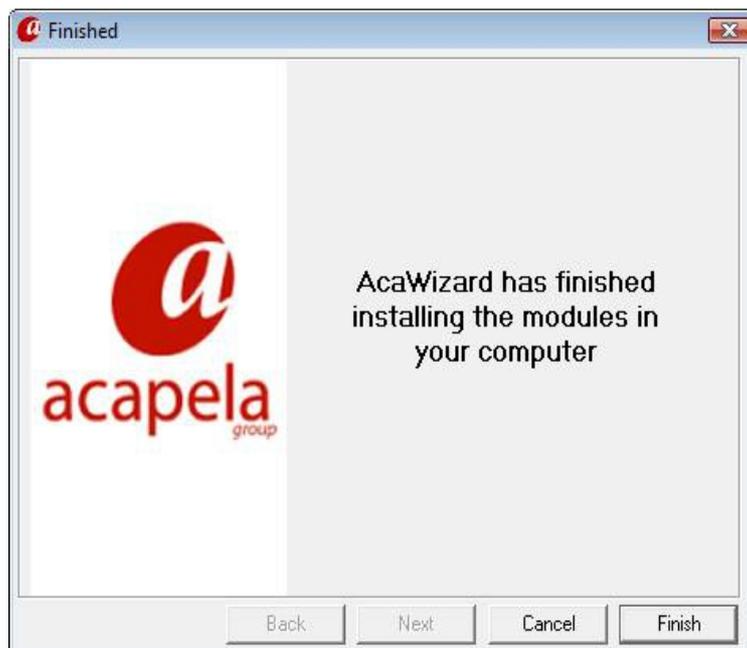


As you see in the legend: items listed in black are already installed; items listed in blue were not installed and will be installed; items listed in red were already installed but will be updated by a newer version. To uninstall one of the items, simply, uncheck the corresponding box.

Click Next for beginning installation:

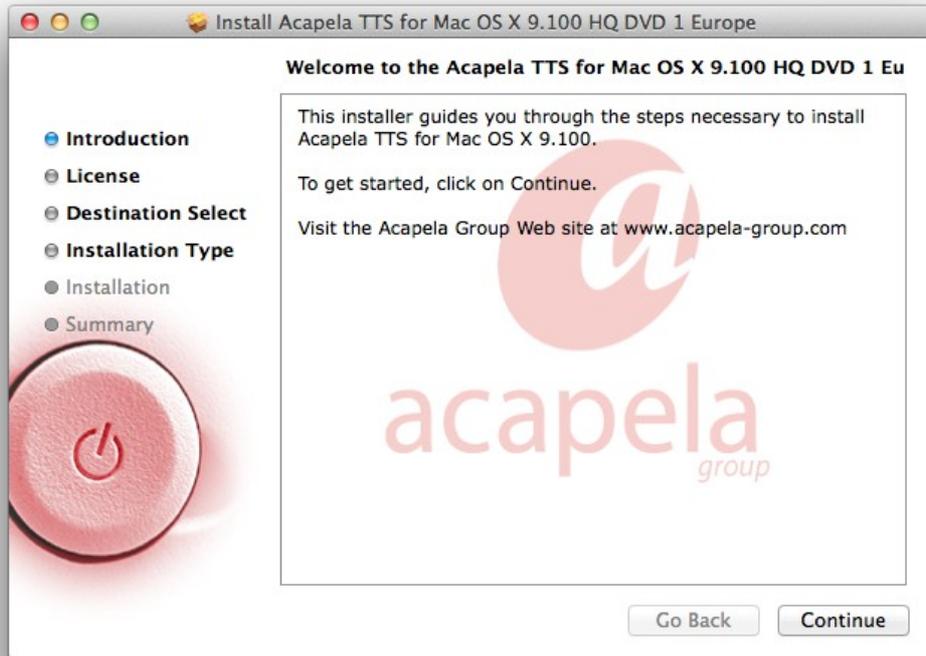


Once the installation is finished, click Finish

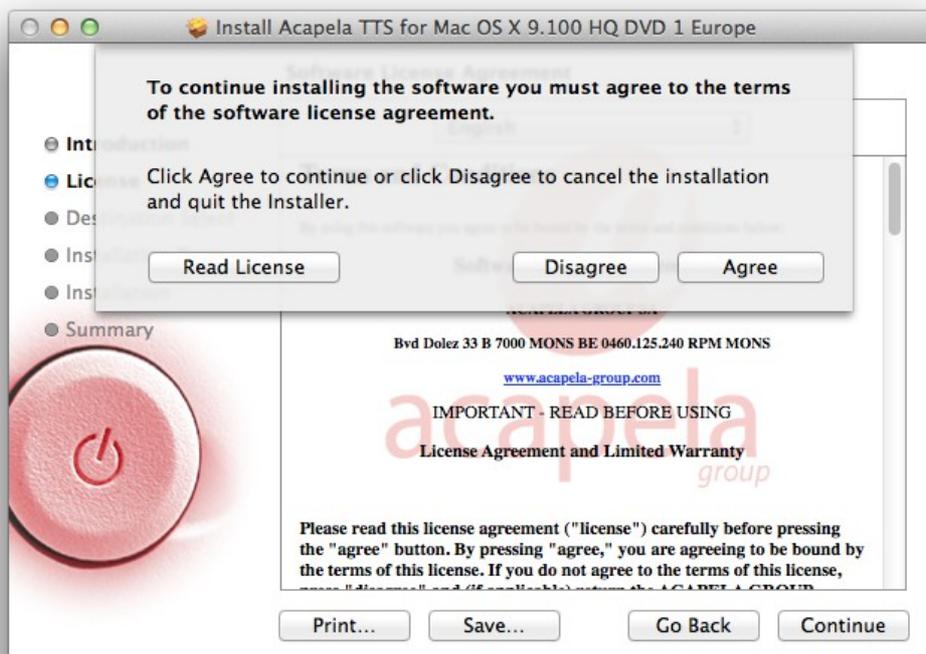


### 2.2.2. Mac OS X

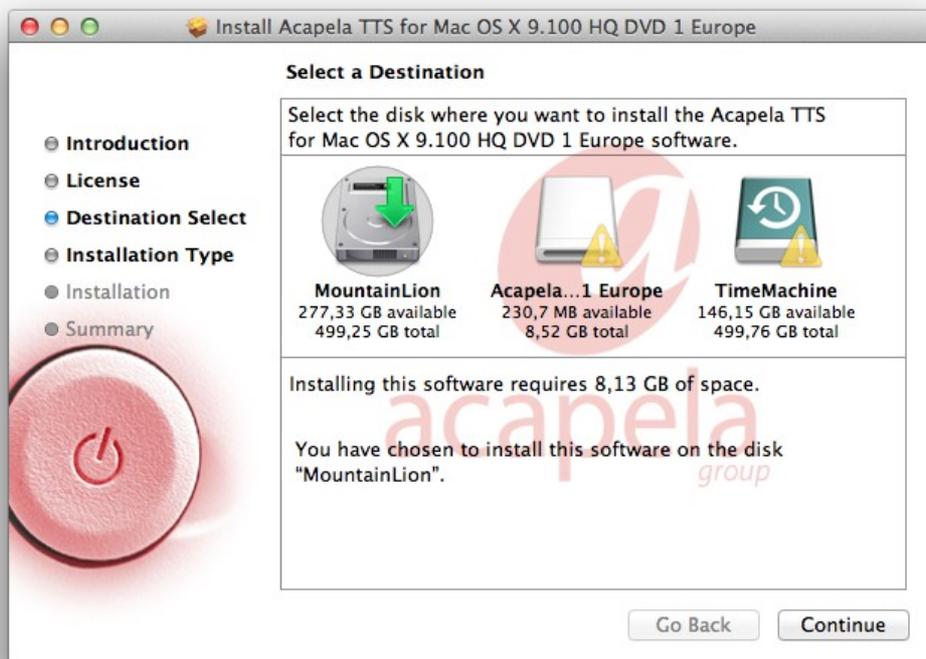
The installer is packaged as a `.dmg` file. Please note that from version 8.100 and higher, there are several `.dmg` files, for example one for Europe and one for the rest of the world. Open the file to mount the image disk and double-click on the `.mpkg` file to start the installer. The welcome screen dialog will pop up. Press the "Continue" button.



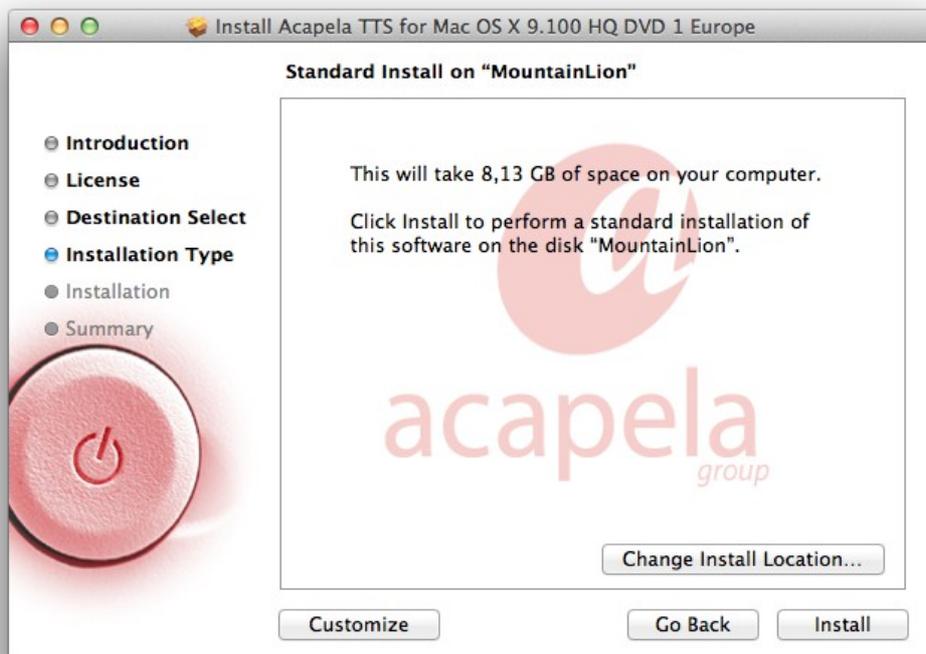
Carefully read the license agreement and, if you agree, select the "Agree" button and then click the "Continue" button.



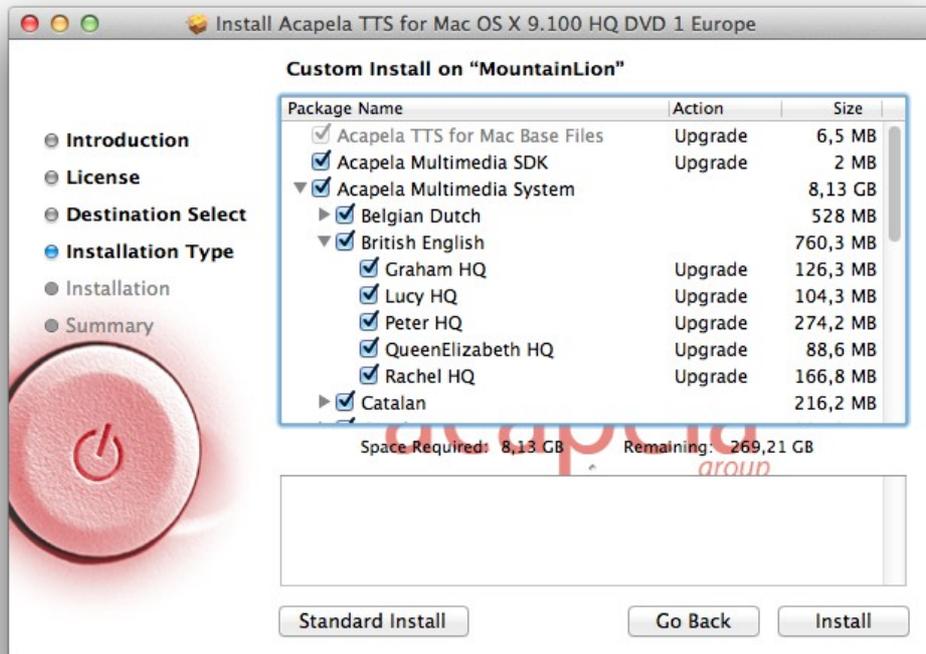
Now you will be prompted with the following screen which propose you to choose the disk where you will install Acapela TTS for Mac. Please note, however, that your choice is restricted to the system disk (the one on which Mac OS X you have booted on is installed). Additionally, this screen is skipped but you can come back to it using the "Change Install Location..." button on the following screen.



You are now presented the pre-installation summary.



Click on the "Customize" button to choose which components of Acapela TTS will be installed.



Here you can select or unselect individually most of the packages, especially the different voices.

Acapela TTS will be installed on the system disk, you cannot choose another disk as it will be installed in the `/Applications` directory. When you are ready you can click on the "Install" button to complete the installation process.

### 2.2.3. Linux-based Systems

There is no installer for Linux. The whole distribution is packaged as single big compressed archive.

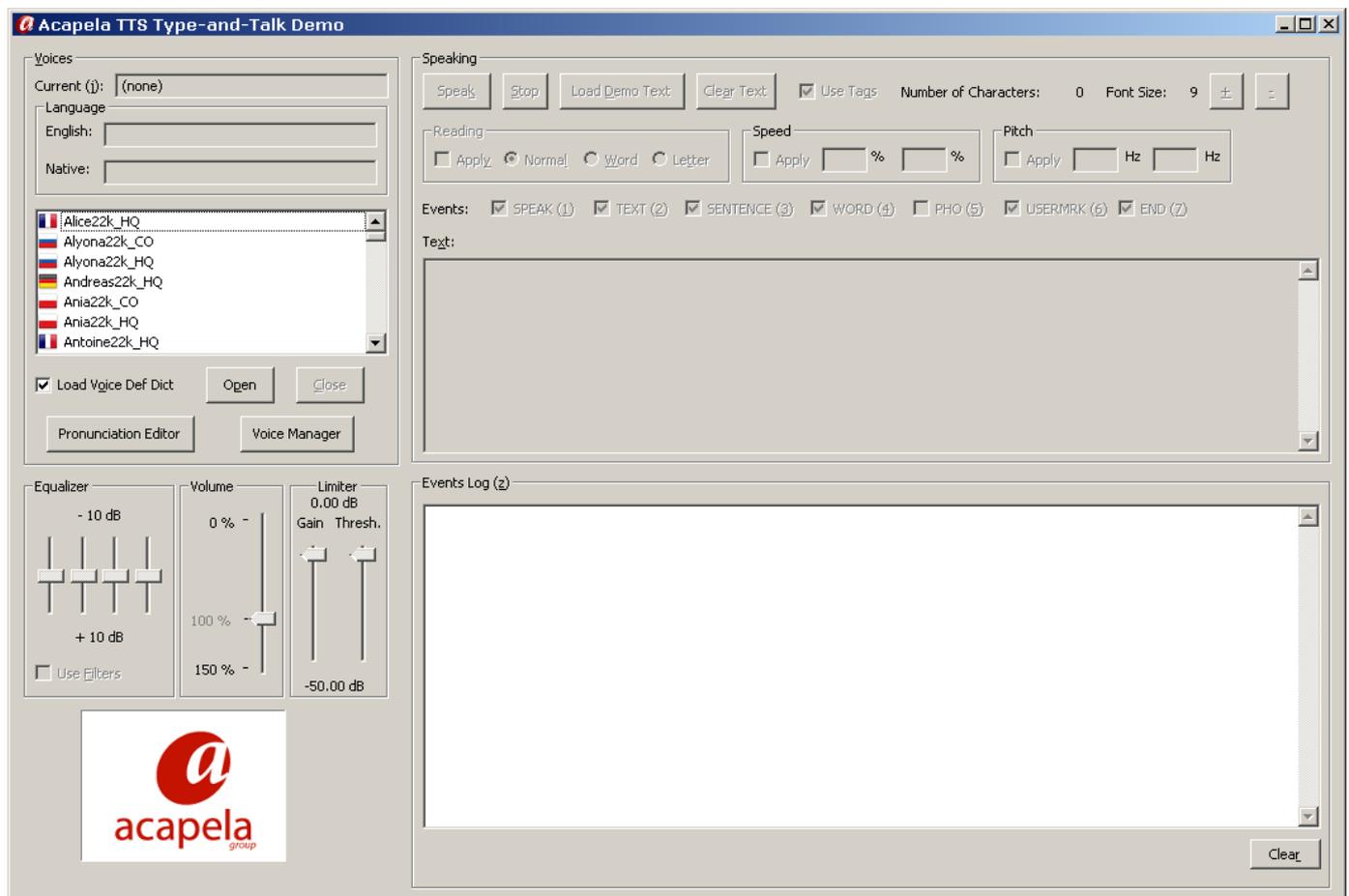
## 3. Testing the TTS Engine

This chapter presents a simple way to check that the package has been correctly installed and explains how to quickly have a demonstration of our text to speech technologies.

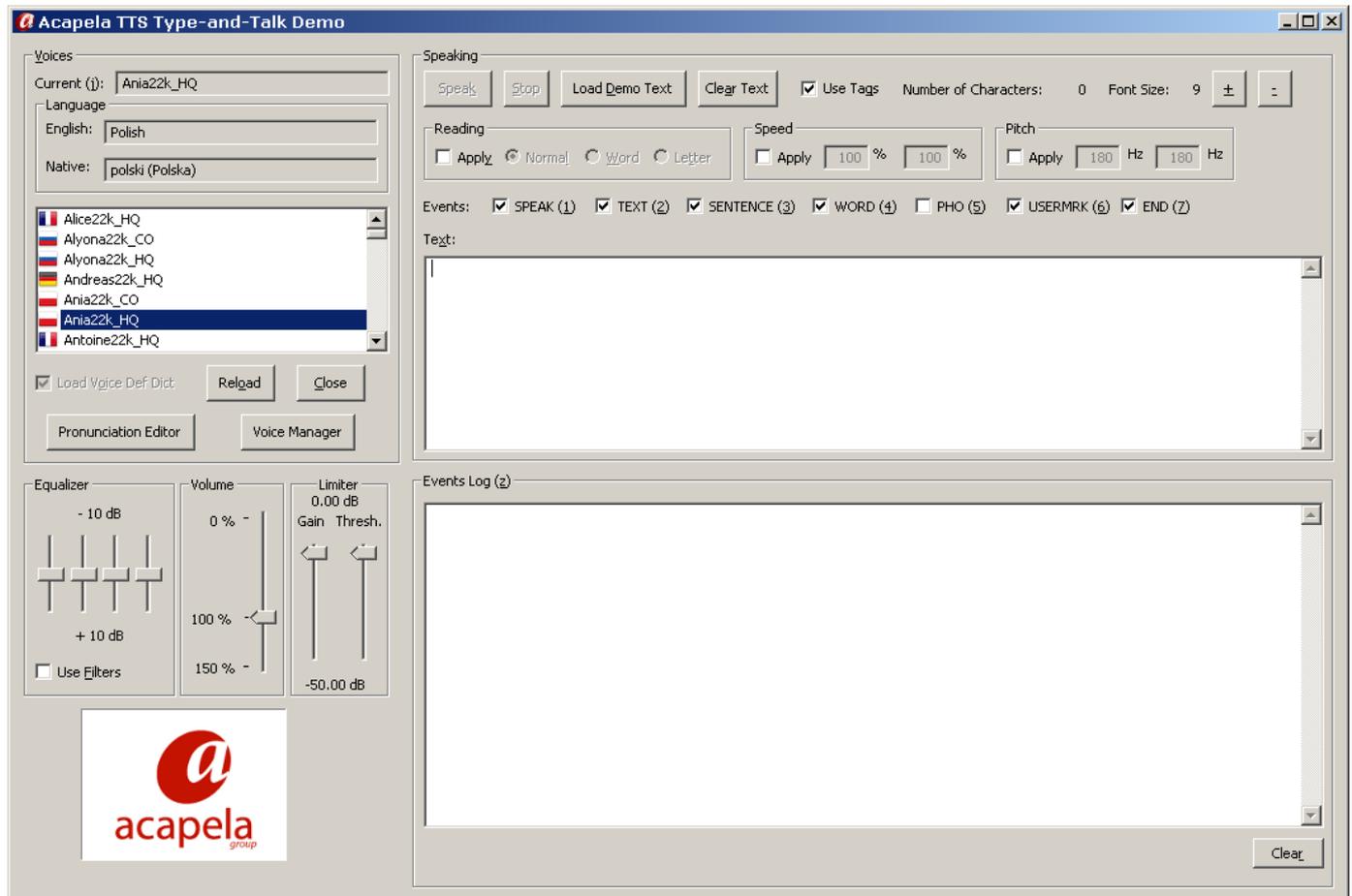
### 3.1. Testing the TTS Engine on Windows

Launch `AcaTTSDemo.exe` from the start menu (under Start → All programs → Acapela Group → Acapela TTS 8.200 for Windows → bin → TTS Demonstrator).

If AcaTTS is correctly installed, the demonstration software AcaTTS Demo appears on your screen. Select a voice (and thus a language) in the dialog box.

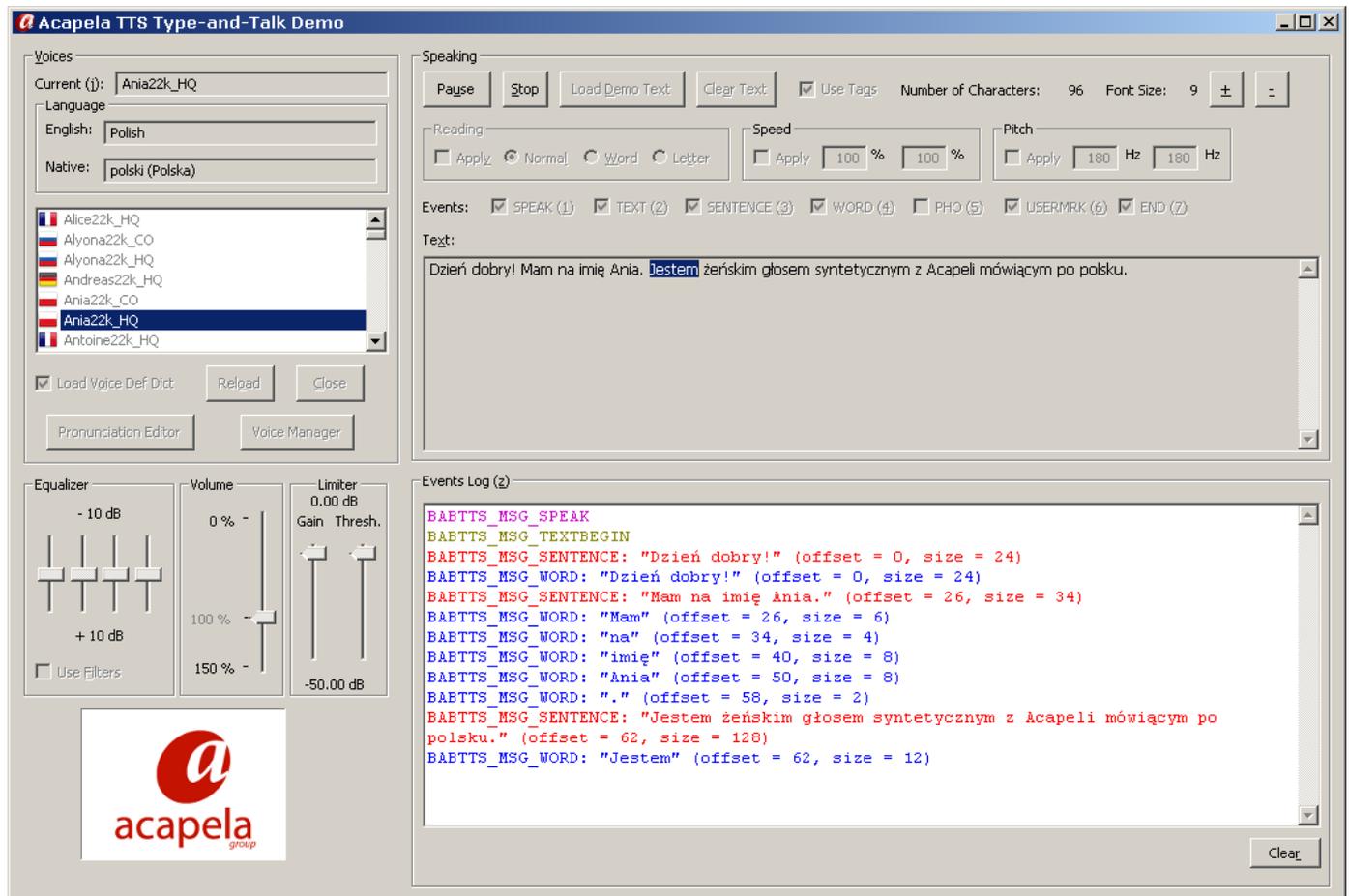


Select a voice and click on the "Open" button. Then type a text and click on the "Speak" button.



You hear the synthesized text on the audio output (headphones or speakers). The word synchronization is shown on the text with a blue highlighting.

You may also see the different events that AcaTTS generates in the "Events Log" window.



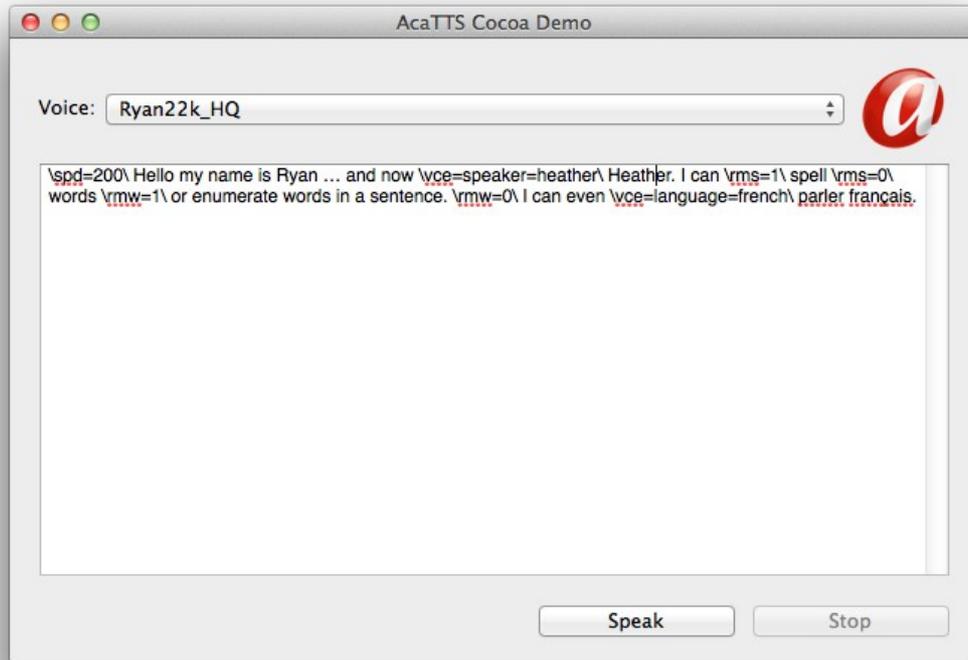
Press the "Stop" button to halt the audio output.

It is possible to stop temporarily the speech by pressing the "Pause" button which will morph into the "Resume" button. Clicking on the "Resume" button morph it back into the "Pause" button and will resume speaking the text.

## 3.2. Testing the TTS Engine on Mac OS X

### 3.2.1. Testing with AcaCocoaDemo

To have a quick demonstration of our voices, you can use `AcaCocoaDemo` located in `/Applications/Acapela TTS for Mac/`.



Choose a voice, type your text in the main window and press the "Speak" button.

## 4. Text Tags

Acapela TTS uses the "Microsoft SAPI 4 Tag Pattern" to read at run-time some specific information in the text.

### 4.1. Exceptions



The only exception is if you are using the SAPI 5 layers. If you do so, you will have to use XML Tags. Please refer to the document "Supported XML SAPI 5 tags" ([ACA MUL-SAPI5](#)) for more information on those.

### 4.2. Text Tags Syntax

Tags are constrained by these rules of syntax:

- A tag begins and ends with a backslash character (\).
- It's case-insensitive. For example `\RST\` is the same as `\rst\`, `\rSt\`, `\rsT\`, `\RsT\`, etc.
- Spacing has effect. For example, `\rst\` is not the same as `\ rst \`. The former won't be recognized.
- To include a backslash character in tagged text
  - Outside a tag: use a double backslash (`\\`). If user wishes to speak a file name as `C:\Windows\System32\test.txt`, then it should double up the backslashes (i.e. `C:\\Windows\\System32\\test.txt`).
  - Inside a tag: put the whole chunk containing backslashes between double quotes. For example: `\audio=play="C:\My Wav Files\some_sound.wav"\`.
- When the engine encounters a tag it does not understand, the tag is simply ignored (not read).
- Some tags are persistent from one text to another. For example, if the `\spd=150\` is present in a text, all following texts will be spoken with a speed of 150 words per minute, until a `\rst` or another `spd` tag is encountered or another voice is selected.

### 4.3. AcaTTS Supported Tags

Here is the list of tags supported in AcaTTS. All these tags are correctly processed for a HD voice (MBROLA technology). Due to limits inherent to the unit selection technology, some tags aren't supported for HQ voices, suitable only for HD voices.

#### 4.3.1. `\col=wgt=val1,val2[,val3,[val4...]]\`

Colibri-specific tag. It allows the morphing between the base voice and its first variant (`val2`). If there's more than one variant embedded in the voice data, each variant has a `valN` where `N` is the number of the corresponding variant. Otherwise, if there's no variant available, only the base voice, this has no effect.

### 4.3.2. `\emph\`

---

This tag sets emphasis on the next word. Emphasis may be rendered differently by different voices/languages. A word can be explicitly de-emphasized by using the `\emph=0\` tag.

### 4.3.3. `\equ=preset\`

---



Change the equalizer preset at runtime. You can manage the list of presets with the Voice Manager, see [Chapter 7](#) for more information about equalizer presets.

### 4.3.4. `\equ=val1;val2;val3;val4\`

---

Where `val1`, `val2`, `val3`, `val4` are in a range from -100 to 100 and affect the frequency bands of 275Hz, 2.2kHz, 5kHz and 8.3kHz, respectively.

### 4.3.5. `\ls=number\`

---

`number` is the minimal (inclusive) threshold required to consider a block of consecutive *symbols* or *letters* to be pronounced as a group of symbols instead of being pronounced individually. The special values of 0 and 1 are *not supported*.

Example: In the text, `I want to aggregate these *** and yyy, \ls=2\like this *** and yyy.`, the first part is pronounced like "asterisk asterisk asterisk and y y y" while the second part sounds like "group of asterisk and group of y".

#### Note

A good way to cancel its effect in a text, instead of using `\rst\`, is to reuse it with a large integer.

This tag persists on the next texts being spoken until the next voice change or the next tag cancelling it out.

### 4.3.6. `\mrk=number\`

---

This tag indicates a user bookmark in the text.

When AcaTTS encounters this tag, it notifies the user's application (see chapter "Using the Callback Mechanism" in the Developer Guide — [ACAMUL-SDK](#)). `\mrk=0\` is reserved. This tag has no use in the AcaDemo application.

### 4.3.7. `\skipaudio=number\`

---

This tag allows to select parts of texts to mute. When `number` is 1 or 2, the text following the tag is muted, when `number` is 0, the sound is sent normally to audio output. In contrast to 1, option 2 ignores too the callback events.

Option 1 is then usefull if we want to analyse the full text or to force the pronunciation of a part of a text in a specific context.

Example: "This part will be read, `\skipaudio=2\`while this one will not, `\skipaudio=0\`simple, right?"

### 4.3.8. `\pau=number\`

This tag inserts a pause of the specified number of milliseconds in the speech.

Example: "I am `\pau=2000\` ready." → a pause of 2 seconds is introduced after the word "am". The algorithm is limited to pauses of 5 seconds.

#### Note

When a punctuation sign (. ! ? : ; , ( ) [ ] { } ' ' ` " ) is preceding by a pause tag, the silent time of these punctuation is overridden by the milliseconds given by the tag.

Example: in "Goodmorning`\pau=100\`. Goodbye." the sentences are separated by a silence of 100 milliseconds instead of default silence. It allows to define precisely the length of the silence associated with a punctuation sign without affecting the overall prosody.

### 4.3.9. `\paumode=number\`

For intelligibility purposes, pauses are automatically inserted by AcaTTS when playing the text, i.e. for sequences of space-separated numbers like `46 27, 108 95`. This feature can be turned off by with `\paumode=1\` and turned (back) on with `\paumode=0\`.

Example: `\paumode=0\ 1 2 3. \paumode=1\ 1 2 3.` The first sequence will be pronounced more slowly (because of the automatic pauses inserted between 1, 2 and 3) than the second one.

This tag persists on the next texts being spoken until the next voice change or the next tag cancelling it out.

### 4.3.10. `\pit=number\`

This tag sets the baseline pitch of the voice to the specified value in Hertz. The actual pitch fluctuates above and below this baseline following the prosodic rules. number must be in the range from 50 (%) up to 200 (% of the base pitch), about 180-190 Hz for a woman and 100-110 Hz for a man, depending on the voice, being the default.

For example, `\pit=70\` will have no effect on a female HD voice, but will work on a male voice.

#### Note

Not processed with HQ voices, unless using `pit2vct` (see [ACAMUL-SDK](#)).

This tag persists on the next texts being spoken until the next voice change or the next tag cancelling it out.

### 4.3.11. `\prn="phonetic string"\`

With this tag, a user can synthesize a specific pronunciation inside a text. The phonetic string is composed of phonemes followed by space characters (phonetic alphabet is language-dependent). This tag is only suitable for inserting single words into the text. Unpredictable errors (mainly prosody) can occur when inserting greater units.

Example: I will say: `\prn=h e l @U1\.` is equivalent to "I will say: hello.."

#### Notes

This tag has not the same meaning in AcaTTS than in the standard SAPI norm, it does not apply the same pronunciation for future words.

The `\prx\` tag (see below) is preferred over this one.

#### 4.3.12. `\prx="phonetic string"\`

This is the same as the `\prn\` tag but is the preferred form.

#### 4.3.13. `\prx=%1nature%word\`

This tag allows us to fix the `nature` of a `word` in a sentence. This can be relevant to remove a potential ambiguity between identical words pronounced differently. `nature` can be chosen in the set {NOUN, ADJ, VERB, ADV, PARTPASSE, PARTPRES, CHIF, INFINIT}.

Example: "The queen and Alice `\Prx=%1VERB%read\` a book." makes sure 'read' will not be pronounced in its pas participle form.

#### 4.3.14. `\rms=number\`

Sets the reading mode to spelling out each letter of each word (`number` is 1), or turns it off (`number` is 0).

#### 4.3.15. `\rmw=number\`

Sets the reading mode to leaving audible pauses between each word (`number` is 1), or turns it off (`number` is 0).

#### 4.3.16. `\rmu=number\`

When turned on, it allows the synthesizer to read intelligibly series of attached words as "CFMutableString". Each word must start with a capital letter, then the word will be read like a sequence of slitted words : "CF Mutable String".

To turn on the feature the number is 1, and 0 to turn it off.

This tag persists on the next texts being spoken until the next voice change or the next tag cancelling it out.

#### 4.3.17. `\rpit=number\`

This tag sets the relative pitch. `number` must be in the range from 50 up to 200, 100 being the default pitch of the voice.

This tag persists on the next texts being spoken until the next voice change or the next tag cancelling it out.

#### 4.3.18. `\rspd=number\`

This tag sets the relative speed. 100 is the default speed (about 180 words per minute, depending of the voice). Call `\rst\` to reset to the default speed.

This tag persists on the next texts being spoken until the next voice change or the next tag cancelling

it out.

#### 4.3.19. `\rst\`

This tag resets the engine to the default settings for the current mode.

**Note:** This tag does not affect settings changed using `BabTTS_SetSettings()`.

#### 4.3.20. `\sel=altN\`

Selector-specific tag. It gives an alternative synthesis for the following word. To further explore alternatives, `\sel=altN\` gives the  $N$ -th acoustic alternative for the following word (and thus, potentially, the whole breath group).

Example: The sentence, `I don't like the sound of this \sel=alt3\word.`, takes the third best alternative of the pronunciation of word *word* and *regenerates* the whole sentence.

#### 4.3.21. `\spd=number\`

This tag sets the baseline average talking speed of the voice to the specified number of words per minute. Each voice has a default speed (about 180 words per minute, depending of the voice). Call `\rst\` to reset to the default speed. Range: 1/3 to 3 times the default speed (typically `\spd=60\` to `\spd=539\`).

This tag persists on the next texts being spoken until the next voice change or the next tag cancelling it out.

#### 4.3.22. `\vce=key=value\`

Instructs the engine to change the speaking voice, according to the specified characteristics. The pitch, speed, volume, etc. revert to the defaults for the new voice.

`key` is the name of the characteristic and can be any of the following:

- `language=language`. Requests the engine speak in the specified language (`language` can be Arabic, BelgianDutch, Brazilian, British, CanadianFrench, Catalan, Czech, Danish, Dutch, FinlandSwedish, Finnish, French, German, Greek, IndianEnglish, Italian, Norwegian, Portuguese, Russian, Spanish, Swedish, Turkish, USEnglish, USSpanish). Example: `\vce=language=Spanish\`.
- `speaker=speakername`. Specifies the speaker value of the voice. Beware that the speaker name is different from the voice name. `Heather22k_HQ`, `Heather8k_HQ` and `Heather22k_HM` are voice names. `Heather` is the speaker name. Example: `\vce=speaker=Ryan\`.
- `gender=gendername`. Uses to specify the gender, *male* or *female*, of the actual language to be used (if voice available). Example launched with Heather (required: Heather and Ryan USEnglish voices installed): "Currently Heather is talking, `\vce=gender=male\` and now, it's Ryan."

This tag persists on the next texts being spoken until the next voice change or the next tag cancelling it out.

#### Note

The `vce` tag only works for voices with the same sampling rate. It cannot be used for changing voices designed for different sampling rates.

### 4.3.23. `\vct=number\`

---

This tag controls the Voice Shaping of the voice, please refer to the Developer's Guide for more information (min: 50%-150%).

This tag persists on the next texts being spoken until the next voice change or the next tag cancelling it out.

### 4.3.24. `\vol=number\`

---

This tag sets the output volume. It is formatted as `\vol=volume\` where `volume` is a value in the range 0 to 65535, inclusive. The default value is 65535.

This tag persists on the next texts being spoken until the next voice change or the next tag cancelling it out.

### 4.3.25. `\wrp=number\`

---

Colibri-specific tag. It controls the warping of the voice (min: 50%-150%) and can be used to adjust the tone of a voice.

This tag persists on the next texts being spoken until the next voice change or the next tag cancelling it out.

## 4.4. Audio Tag

---

### 4.4.1. Overview

---

Allows for mixing and configuring background sound/music or inserting sound in the text.

### 4.4.2. Sound files compliance

---

Only **raw PCM 16-bit mono files with the same frequency as the current voice** are supported by the audio tag.

### 4.4.3. Syntax

---

```
\audio=command1[=argument1[;argument2...]]\
```

We can place multiple commands in the same tag (separated by ";"), but if commands are wrongly associated (e.g.: mix and play) the whole audio tag is ignored.

If backslashes (`\`) are to be used in the path, i.e. to specify a file path (for example, on Windows systems, but can also be used on Mac OS X and Linux systems), you must enclose the affected character string (file path) in double quotes (`"`). For example:

```
\audio=mix="c:\mozart.pcm";duration=5000\.
```

Forward slashes (`/`) can also be used within file paths, even on Windows systems, not needing any special character treatment as it is not a character with special semantics for the synthesizer (i.e., `\audio=mix=c:/mozart.pcm;duration=5000\`).

Spaces are authorized as well the same way (i.e., `\audio=mix="c:\moz art.pcm";duration=5000\` or `\audio=mix=c:/moz art.pcm;duration=5000\`).

#### 4.4.4. Commands

---

##### 4.4.4.1. `\audio=mix="filepath"\`

---

**Plays the file in the background**, the speech synthesis will continue during the playing. Following commands (e.g.: stop, pause, resume...) will be applied to this tag until the next Mix or Play=file.xxx tag.

Example: `\audio=mix="c:\mozart.pcm"\I speak with Mozart playing in the background!`

#### Note

All mix commands are enqueued. As soon as a mix command ends, the next one in the queue is then played (Always in background).

##### 4.4.4.2. `\audio=offset=X\`

---

**Skips X milliseconds** at the beginning of the sound.

Example: `We can start again where we left off!`  
`\audio=play="c:\mozart.pcm";offset=5000\.`

##### 4.4.4.3. `\audio=pause\`, `\audio=resume\` and `\audio=stop\`

---

**Pauses, resumes or stops background playing** (mix mode only).

Example: `\audio=mix="c:\mozart.pcm"\ I put the background music on pause!`  
`\audio=pause\ Then I resume it! \audio=resume\ Finally, I stop it! \audio=stop\ It's finished..`

##### 4.4.4.4. `\audio=play\`

---

If the file name is omitted and there is any `\audio=mix\` command enqueued, it **turns the background playing into foreground playing, thus going from asynchronous mode to synchronous mode**. For example, at the end of a text it will cause the foreground music to continue until the end of the sound currently playing rather than immediately stopping.

Example: `\audio=mix="c:\mozart.pcm"\I speak with Mozart playing in the background. Now I'll be quiet and let the music track play until it's finished... \audio=play\.`

##### 4.4.4.5. `\audio=play="filepath" [;argument1[;argument2...]]\`

---

**Plays a sound in the foreground** (synchronous mode).

Example: `Please applaud! \audio=play="c:\bravo.pcm"\ Thank you!`

Optionally, arguments can be given, here are they:

##### 4.4.4.5.1. `continue`

---

**Makes a sound continue in the background** (asynchronous mode). There must a `duration=timeduration` or `until=timeposition` arguments, it **turns the foreground playing into background playing (thus going from synchronous playing to asynchronous mode)** when the limit specified by those other arguments is reached.

Example: `\audio=play="c:\bravo.pcm";duration=100;continue\ Thank you very much!`. It will play `bravo.pcm` in the foreground for 100 milliseconds then continue playing it in the background while saying "Thank you very much!".

#### 4.4.4.5.2. `duration=timeduration`

**Plays the sound for `timeduration` milliseconds** (play or mix mode) and then stop reading it. Examples:

- `Now you will listen to Mozart for five seconds`  
`\audio=play="c:\mozart.pcm";duration=5000\`
- `\audio=mix="c:\mozart.pcm";duration=1250\The music will stop while I speak...`
- `\audio=mix="c:\mozart.pcm"\ I'll let the music play for five seconds after speaking... \audio=play;duration=5000\`

#### Notes

The `until` and `duration` commands really are different. `until` is absolute in play time specification in respect to the beginning of the sound whilst `duration` is relative in play time specification in respect to it, `duration` is thus relative (and sensible) to the amount of sound already played.

As a notable property, if no `offset` argument is specified, which implies no offset from the beginning of the sound when playing it, and no previous `\audio=mix\` or `\audio=play\` command was specified, `until` and `duration` arguments are confounded, they have the same effect.

#### 4.4.4.5.3. `until=timeposition`

**Plays the sound until the position `timeposition` milliseconds within the sound is reached** (play or mix commands) and then stops reading it. Examples:

- `Now you will listen to Mozart for five seconds`  
`\audio=play="c:\mozart.pcm";until=5000\`. See the notes.
- `\audio=play="mozart.pcm";duration=2000\ The music will stop when I start speaking, it lasts only two seconds, then, when I am finished speaking, it will continue playing the music where it stopped for three seconds.`  
`\audio=play="mozart.pcm";offset=2000;until=5000\`

#### Notes

The `until` and `duration` commands really are different. `until` is absolute in play time specification in respect to the beginning of the sound whilst `duration` is relative in play time specification in respect to it, `duration` is thus relative (and sensible) to the amount of sound already played.

As a notable property, if no `offset` argument is specified, which implies no offset from the beginning of the sound when playing it, and no previous `\audio=mix\` or `\audio=play\` command was specified, `until` and `duration` arguments are confounded, they have the same effect.

#### 4.4.4.6. `\audio=repeat=status\`

---

When `status` is `on`, continuously **repeats the foreground or background** sound (play or mix command). When it is `off`, does not repeat it (anymore).

Example: `\audio=mix="c:\bravo.pcm";repeat=on\I speak while they applaud! bla, bla, bla, ... bla, bla, bla... \audio=play;repeat=off\`.

If `status` is a positive integral number, possibly zero, it will be the number of times to play the sound from the beginning as soon as its end is reached *in addition to the regular play of the sound*. In the end, the sound is played `status` times + 1.

Example: `\audio=play="c:\gong.pcm";repeat=1\I hear two gongs then this text... bla, bla, bla, ... bla, bla, bla... Beware that despite repeat=1, two gong.pcm will be played two times!`

#### 4.4.4.7. `\audio=volume=percentage\`

---

**Sets the volume** of the sound to `percentage` % of its base level. 100 is the base level. Lesser than 100 reduces the volume, greather than 100 raises it (which can lead to distortion and saturation).

Examples:

- `\audio=mix="c:\mozart.pcm";volume=25\I speak with Mozart playing smoothly!`
- `\audio=mix="c:\mozart.pcm";volume=200\I speak with Mozart playing loudly!`
- `\audio=mix="c:\mozart.pcm"\I will turn down the music slowly...\audio=volume=80\  
\pau=1000\  
\audio=volume=60\  
\pau=1000\  
\audio=volume=40\  
\pau=1000\  
\audio=volume=20\  
\pau=1000\  
\audio=volume=10\  
\pau=500\  
\audio=volume=5\  
\pau=500\  
\audio=volume=0\  
\pau=500\  
\`

### 4.4.5. Limitations

---

- Only absolute paths are accepted as links to sound files, except when using the Audio Resource Directory feature as explained in [ACAMUL-CCD](#).
- Only **raw PCM 16 bits mono files with the same frequency as the current voice** are supported.
- Sometimes you may hear a little crack or an unexpected silence while playing text using this tag. This is only a real-time playback issue; recorded files will not be affected.

## 5. Acapela TTS License

Acapela TTS uses a licensing system that allows a fine control of the resources and provides the stability and reliability that is needed in a professional application. Note that this licensing system is the same for HD and HQ voices.

This licensing scheme is based on a unique *license identity*, which can be provided from different locations (dongle, computers, ...) and a license file describing the type and information of the license associated with the license identity (number of concurrent channels, authorized features, remaining days for evaluation, ...). This concept permits the Acapela TTS system to be compliant with all kind of applications you expect.

During the development process, you will probably use licence files that are "computer" locked while in the final product, you will have to bundle the TTS engine within your products.

In order to get a working license, you will have to use the *License Manager*.

Consult the sales department ([sales@acapela-group.com](mailto:sales@acapela-group.com)) to determine which type of license best suits your needs.

### 5.1. The License Identity

---

#### 5.1.1. Site Authorization

---

With this kind of identity, the License Manager first computes a primary key that will uniquely identify your computer, based on several hardware components.

Once you send your primary key along with your requests, you will receive a license file that will allow the AcaTTS to work on your particular computer. This file will be locked to your computer and will not work with any other computer.

#### 5.1.2. Dongle Authorization

---



For advanced applications (like telecom), the need to have a robust system is a priority. The software site identity is not relevant for such applications, as it implies the creation of a new license file each time the application is moved from one computer to another.

The hardware key (or dongle) is independent of the architecture of the computer you use thanks to a unique identifier stored in this hardware key.

The license file can be easily and safely saved or copied, which means that moving the license to another computer is as simple as copying a file and moving the hardware key between two computers.

Contact us for more information.

#### 5.1.3. Bundling

---

This special licensing scheme is for CDROM distributed applications where the final user has no relation with Acapela Group and/or his local distributors.

### 5.2. The License File

---

The license file is the complement to the license identity and is used to verify the validity of this identity.

This file contains also information about:

- The voices you are allowed to use in your application.
- The number of channels you are allowed to use in your application.
- The features you are allowed to use depending of the type of license agreement you have:
  - Saving the produced speech to files.
  - Retrieving the produced buffers at runtime (for telecom applications).
  - Retrieving generated phonemes info at runtime.
  -  Authorizing the user to work with the SAPI 4 or the SAPI 5 layers.

## 5.3. The License Manager

---

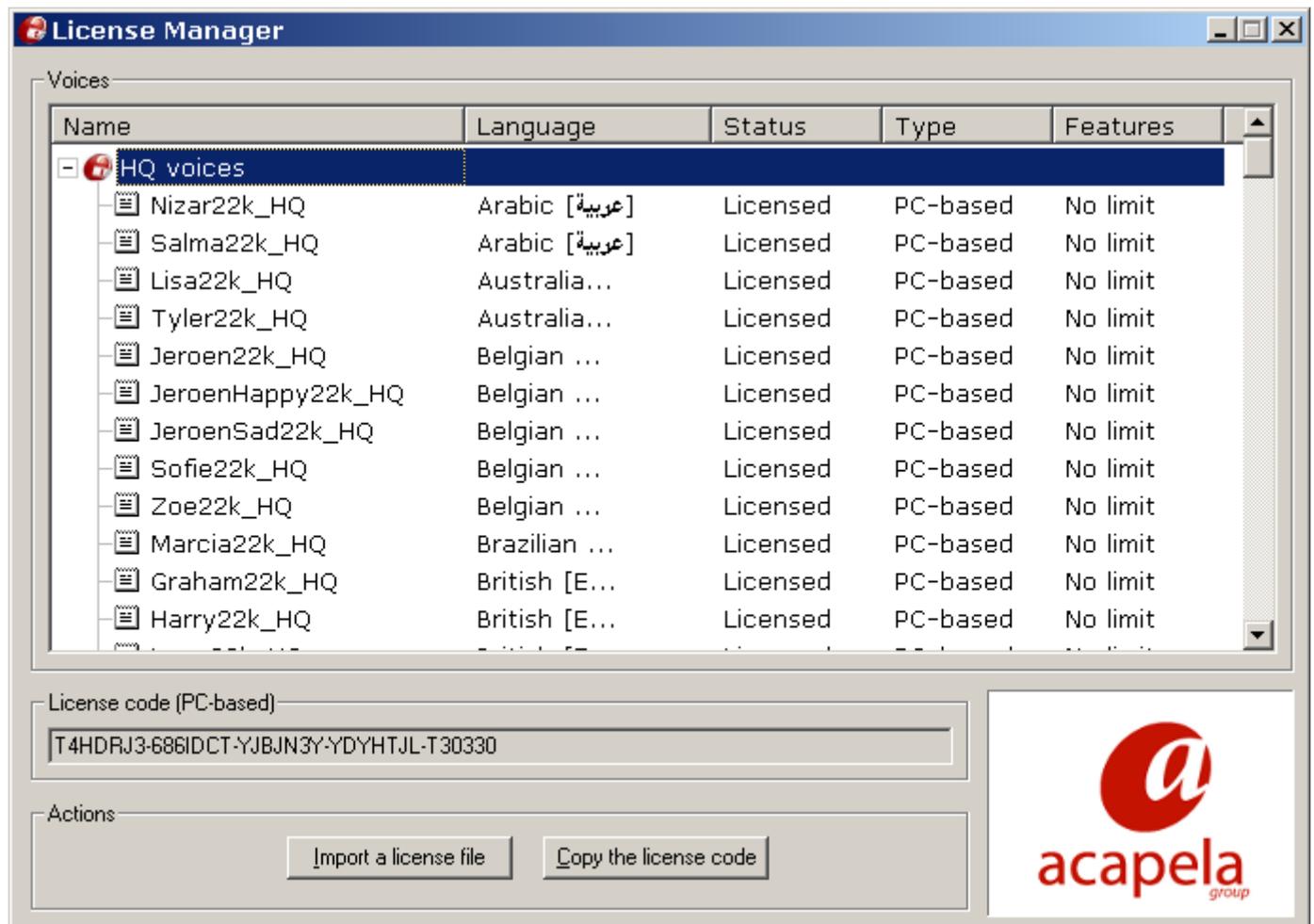
### 5.3.1. Microsoft Windows

---

For Microsoft Windows-based systems, all license-related operations are made through the **License Manager**.

#### **Note**

You need administrator rights to manage licenses.

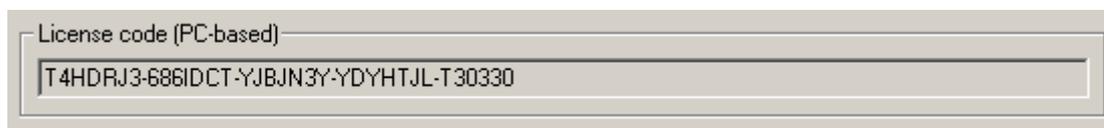


#### 5.3.1.1. Getting a Free 30 Days Trial License

30-days trial licenses aren't available anymore since version 8.100 of Acapela TTS.

#### 5.3.1.2. Getting a Developer License (Computer-based)

- Copy your site code (primary key). You can use the "Copy to Clipboard" button to copy it to your clipboard



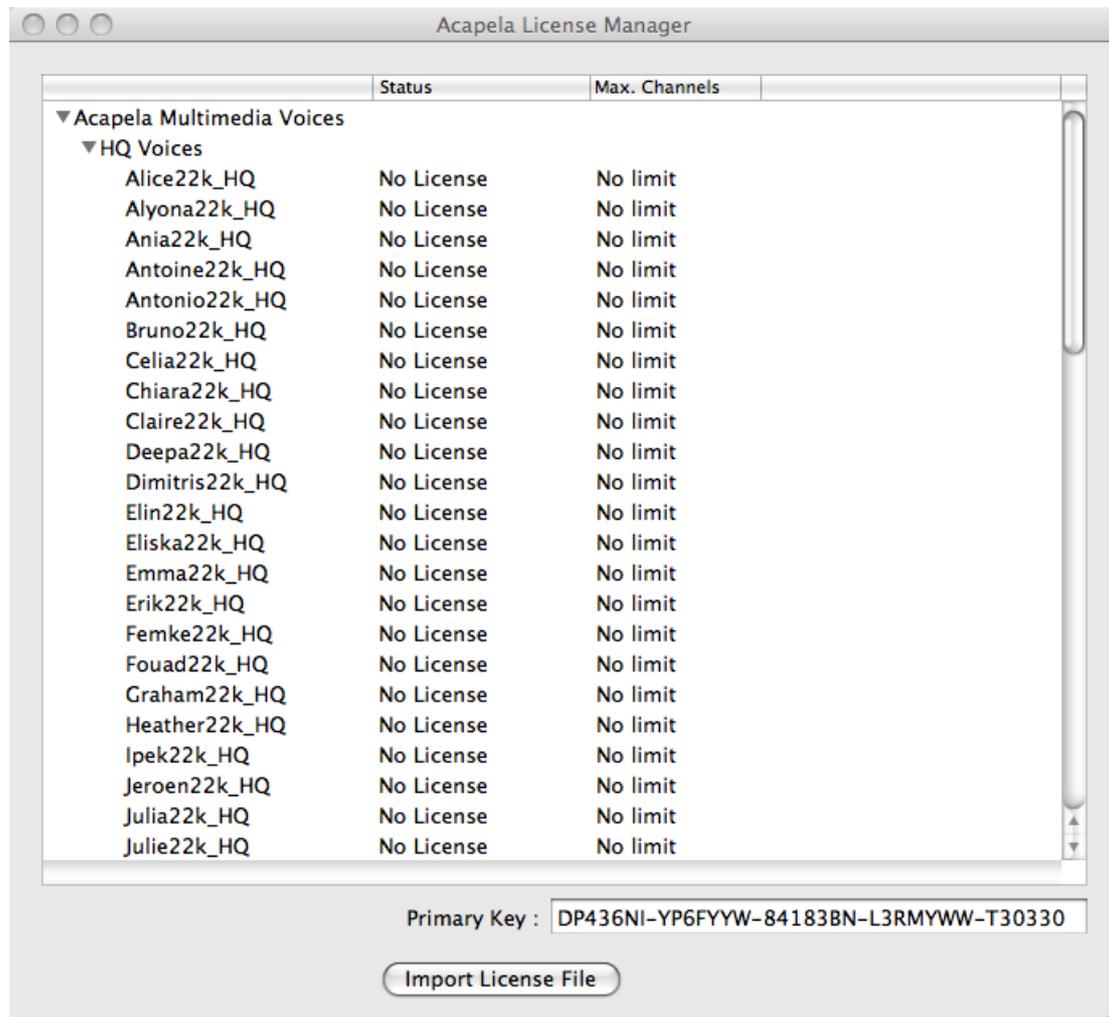
- Ask your contact person at Acapela Group to generate a license file corresponding to your site code or send it to [support@acapela-group.com](mailto:support@acapela-group.com).
- Import the license file: import the license file sent to you into the License Manager using the "Import License" button or, more simply, drag and drop it into the application.

#### 5.3.1.3. Other Operations

There're no more "other operations" available since version 8.100 of Acapela TTS.

### 5.3.2. Mac OS X

All license related operations are made through `LicenseManager.app` which is located in `/Applications/AcapelaTTS/`.



#### 5.3.2.1. Getting a Free 30 Days Trial License

No more 30 days trial license from version 8.100 and higher of Acapela TTS.

#### 5.3.2.2. Getting a Developer License (Computer-based)

Primary Key : DP436NI-YP6FYYW-84183BN-L3RMYWW-T30330

Ask your contact person at Acapela Group to generate a license file corresponding to your primary key or send it to [support@acapela-group.com](mailto:support@acapela-group.com).

Import the license file sent to you into the License Manager using the "import license" button.

#### Note

Do not forget to close and reload your speech-enabled application.

### 5.3.3. Linux-based Systems

---

There is no License Manager on Linux systems.

#### 5.3.3.1. Getting a Commercial License

---

Invoke `GetPK`, situated in the Acapela TTS distribution, for getting the primary key of the computer you want to install the license on.

Ask your sales contact at Acapela Group to generate a license file corresponding to your site code or send it to [support@acapela-group.com](mailto:support@acapela-group.com).

Import the license file by placing it in the directory pointed to by the `Path` entry of the `[License]` section of the `License.conf` file.

## 6. User Lexicons

### 6.1. Definition

When using the synthesizer, you might want to change the way a word is spoken. This happens often when you have names, acronyms, foreign words or other peculiar words that do not follow the default pronunciation rules.

Creating a user lexicon (or user dictionary) gives you the opportunity to set the pronunciation of a word exactly the way you like it.

A user lexicon consists of a list of names (which works as a lookup string) and translations.

For instance, let's suppose that we want the Text-to-Speech Engine to actually speak the sentence "as soon as possible" every time the acronym "ASAP" is encountered in the text. To achieve this, you have to add to the active user lexicon file an entry whose name is "ASAP" and whose translation is "as soon as possible".

The translation string can either be specified as text or as phonetic text, by preceding it with phonetic prefix character (#) and using the phonetic symbols defined for the language in use.

#### Note

*Japanese* (ja-JP) and *Mandarin Chinese* (zh-CN) don't support loading nor manipulating user lexicons.

For simplicity of use, a default lexicon is always created for each voice. This lexicon has the same name as the voice. If there is several account on your computer, each user will have its own lexicons.



These lexicons are installed in

- XP: `C:\Documents and Settings\userName\Application Data\Acapela Group\AcapelaTTS\UserLexicons\`, where `userName` is the name of your user account.
- Vista, 7, 8 and 8.1: `C:\Users\userName\AppData\Roaming\Acapela Group\AcapelaTTS\UserLexicons\`.

and have the name of the voice appended with `.dic` (i.e. `Heather22k_HQ.dic`).

Each lexicon is created the first time a voice is loaded.



These lexicons are installed in `/Users/userName/Library/Application Support/Acapela Group/AcapelaTTS/UserLexicons/`, where `userName` is the (raw) name of your user account, and have the name of the voice appended with `.dic` (i.e. `Heather22k_HQ.dic`).

Each lexicon is created the first time a voice is loaded (except for the user who has installed Acapela TTS where all lexicons were created during the installation).

#### Note

These default lexicons are different of the default user lexicons defined in § Erreur : source de la

référence non trouvée.

## 6.2. Features

- A lexicon is language-dependent, it cannot be used with different languages. But inside the same language, different voices can share the same lexicons.
- Each voice can use as many lexicons as you want. You can define a list of default lexicons with the lexicon editor.
- You can add new entries by several ways :
  - A "sounds-like" method: the translation consists of a plain orthographic word that is pronounced as you want.
  - A "phonetic" input: By preceding the translation with a #, you can insert phoneme codes (using spaces as a separator inserted between them) and expect a fine-tuned pronunciation. See the language-specific manual for more information about the usable phoneme set.
- If you want your user lexicons to always be loaded when starting the engine, you have to define them as "default ones" (see paragraph Erreur : source de la référence non trouvée).
- The default lexicon for the voice is automatically loaded and can be edited at any time using the lexicon editor.



Windows:

- On XP: it's located in `C:\Documents and Settings\userName\Application Data\Acapela Group\AcapelaTTS\UserLexicons\`, where `userName` is the name of your user account.
- Vista, 7, 8 and 8.1: `C:\Users\userName\AppData\Roaming\Acapela Group\AcapelaTTS\UserLexicons\`, where `userName` is the name of your user account.



Mac

`/Users/userName/Library/Application Support/AcapelaTTS/UserLexicons/`, where `userName` is the (raw) name of your user account.

## 6.3. Using The Pronunciation Editor (Windows)

The Pronunciation Editor enables you to manager lexicons and their entries specifically for each language.

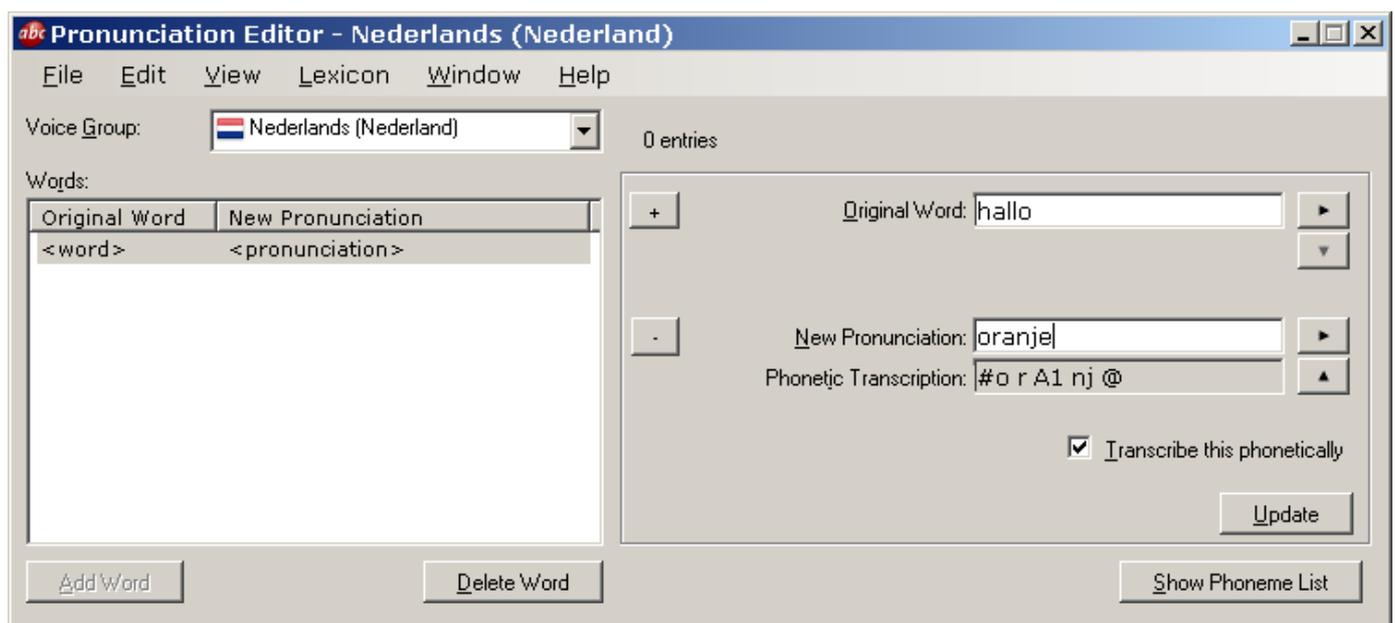
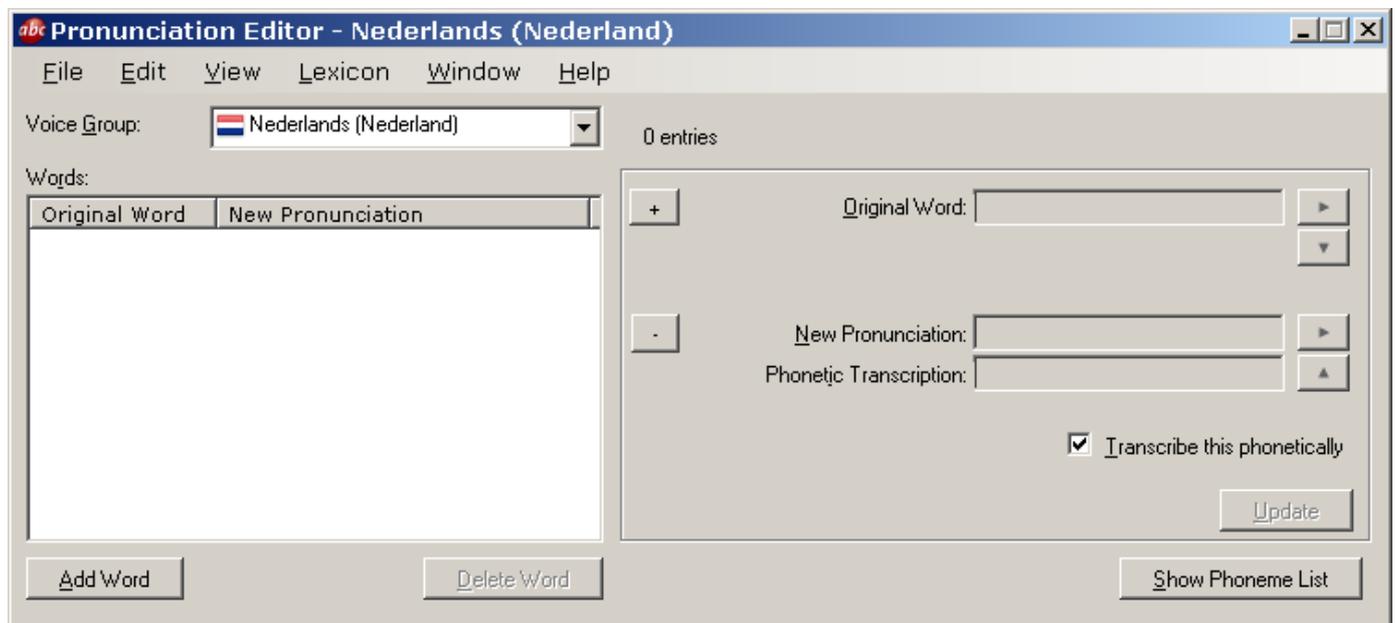
When a given word in a given language has a pronunciation not matching the one you expected, you may "fix" it by entering an entry into the language lexicon. The entry is made up of the original word of which the pronunciation is to be "fixed" and the new pronunciation for that word.

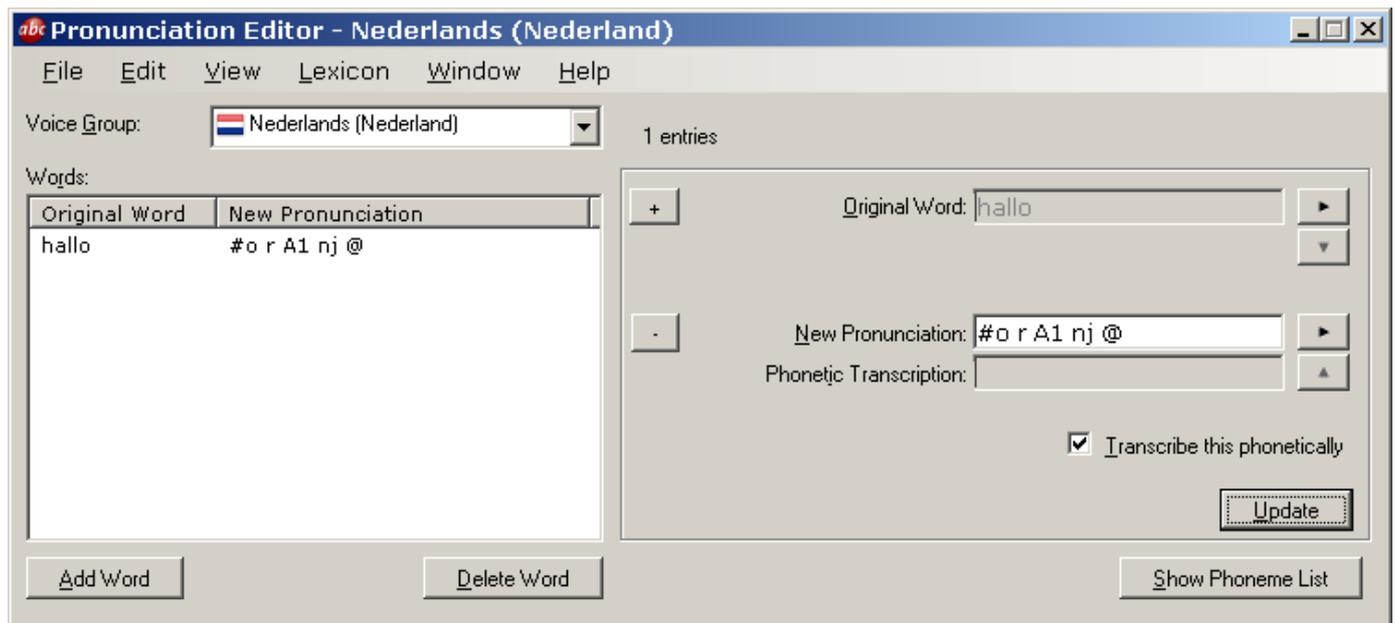
With the Pronunciation Editor, you can:

- add or delete entries in the words list: each entry is composed of a word that's mapped to a new pronunciation
- import and export the current lexicon in AcaTTS or dictman format
- see the list of valid phonemes for the current language
- access directly to the language manual through the Help menu

### Note

*Japanese* (ja-JP) and *Mandarin Chinese* (zh-CN) don't support loading nor manipulating user lexicons. That's why they're not listed in this application.





## 6.4. Using The Java-based Pronunciation Editor (Windows and Mac OS X)

The Pronunciation Editor is a advanced lexicon editor providing the same basic functionality as the Pronunciation Editor does but it also provides additional advanced functionality as you can

- import (massive) lexicons in Elan's Abbreviation/Exception, AcaTTS, LDI, or dictman format
- export the current lexicon in dictman format
- manage default dictionaries for each voice
- manage the currently open dictionaries (in example, close one dictionary or multiple dictionaries) and set which one is active

### Note

*Japanese* (`ja-JP`) and *Mandarin Chinese* (`zh-CN`) don't support loading nor manipulating user lexicons. That's why they're not listed in this application.

### 6.4.1. Location



Mac

LexiconEditor.app is installed in `/Applications/AcapelaTTS/`.

### 6.4.2. Requirements

In order to be run, the Lexicon Editor requires a Java virtual machine to be installed on the machine. Please make sure you have the correct version listed below.

---

## Windows and Linux-based Systems:

---

Java Standard Edition 6.0 32-bit JVM International version by Oracle.

## Mac OS X Systems:

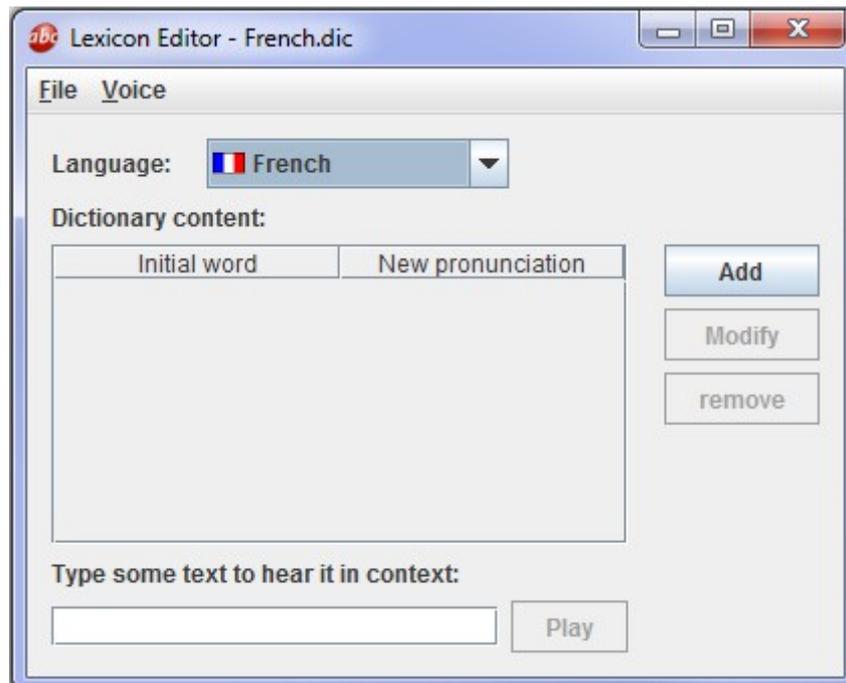
---

Java Standard Edition 6.0 64-bit JVM by Apple (for Mac OS X 10.6) or Oracle's JVM implementation (for Mac OS X 10.7 and beyond).

---

### 6.4.3. Main Window

---



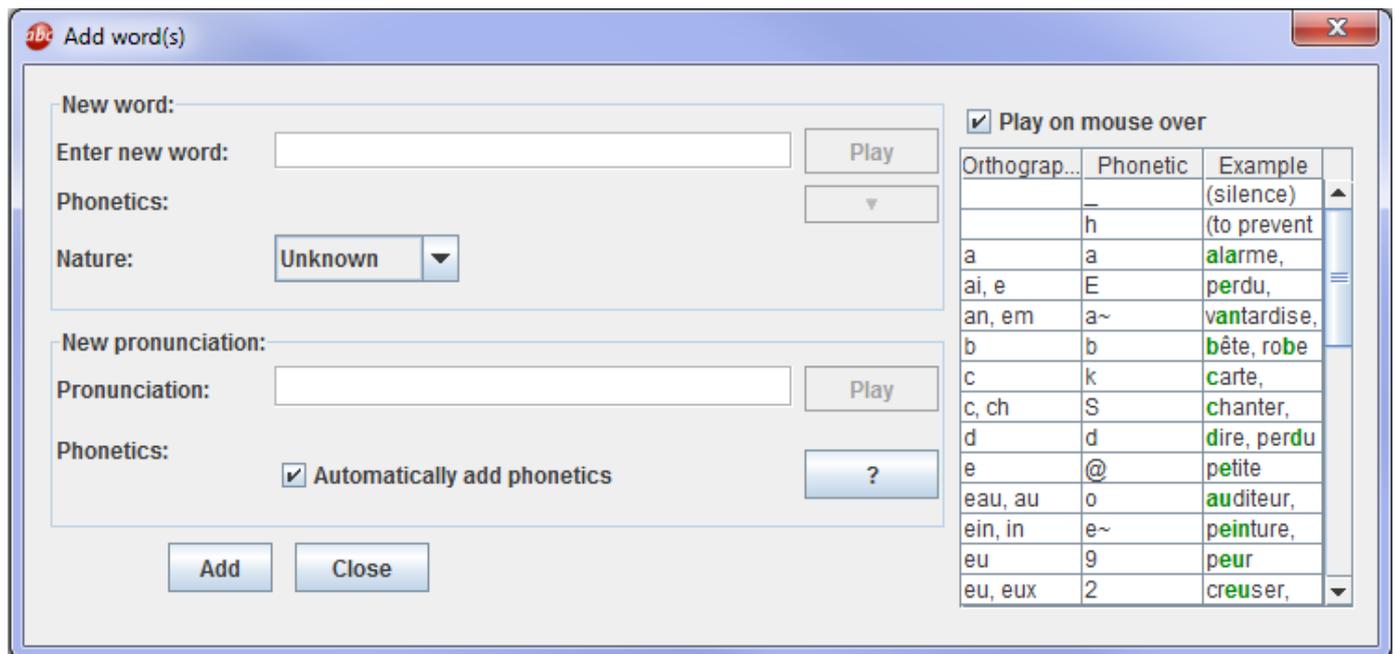
The Lexicon Editor main window allows you to select a language and a voice and manage the content of a lexicon. Using the language list, you can select a language from the ones currently installed on your system. The Main part of the window displays a two-column table listing the entries of the current dictionary. In the left column are the words whose pronunciations are to be changed, in the right column are these words modified pronunciations.

On the right part of the window are the content management buttons:

- the **Add** button adds a new word to the current lexicon
- the **Modify** button changes selected words in the current lexicon
- the **Remove** button removes selected words in the current lexicon.

Finally, the bottom part of the window contains a text field which enables you to test your loaded lexicons by typing some text to hear it and control if it is pronounced correctly.

## 6.4.4. The Add/Modify Window



When clicking on the **Add** or **Modify** buttons from the main window, the Add or Modify window will pop-up. This window enables you to specify the way a word is to be pronounced.

### 6.4.4.1. New Word Frame

The **Enter new word** text field specifies the word whose pronunciation is to be changed. The **Play** button allows you to hear the current word pronunciation. Below the text field will be displayed this word current phonetization. The **down-arrow** button will copy this phonetization to the pronunciation text field below. Finally, the **nature** list is to select this word grammatical nature.

### 6.4.4.2. New Pronunciation Frame

The **pronunciation** text field allows you to set the word new pronunciation. If you prefix your input with a pound sign '#', the input is considered as a phonetic input (a list of space separated phonemes), otherwise it is considered as a literal translation. For phonetic input, a checker is provided and incorrect phonemes will be displayed in red as you type. If you entered a literal translation, the phonetization of your input will be displayed. The **Play** button allows you to hear the new pronunciation.

Using phonetics instead of literal translation is preferred, as literal say as are susceptible to be modified by later TTS engine upgrades. By using phonetics, you ensure your lexicons will remain consistent in time. In order to ease the use of phonetic translations, if the **Automatically add phonetics** check box is selected when adding a literal translation, the lexicon editor will take its phonetization into account instead.

### 6.4.4.3. The Phonetic Table

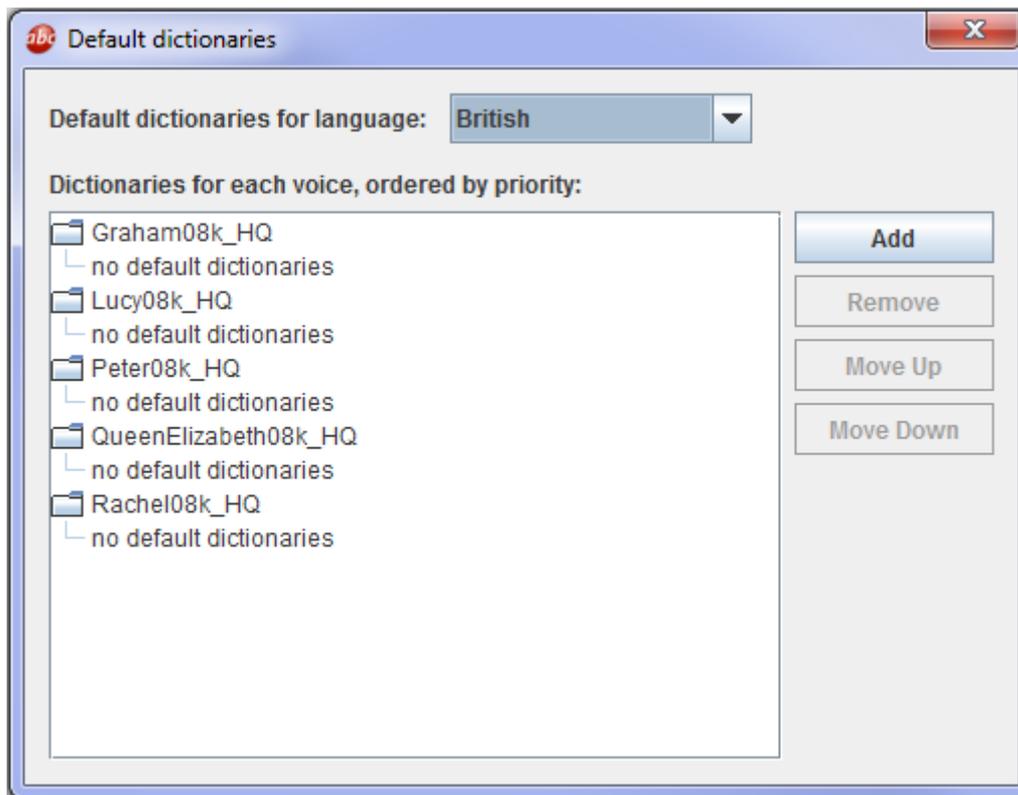
On the right part of the window is a table that lists all the phonemes available for the current language. If the **Play on mouse over** check box is checked, you will hear an example pronounced when you keep the mouse cursor over a line.

The phoneme table is organized in three columns, the first one is an orthographic representation of the phoneme, the second one is the phonetic symbol used, the last one displays examples of words that use this phoneme.

When Adding a new word to a lexicon, the **Add** button inserts a word to the current dictionary and reset the window interface to enable several entries to be inserted in a row.  
When Modifying words, you can modify multiple words: the **Modify** button updates the current entry and jumps to the next selected word (if any).

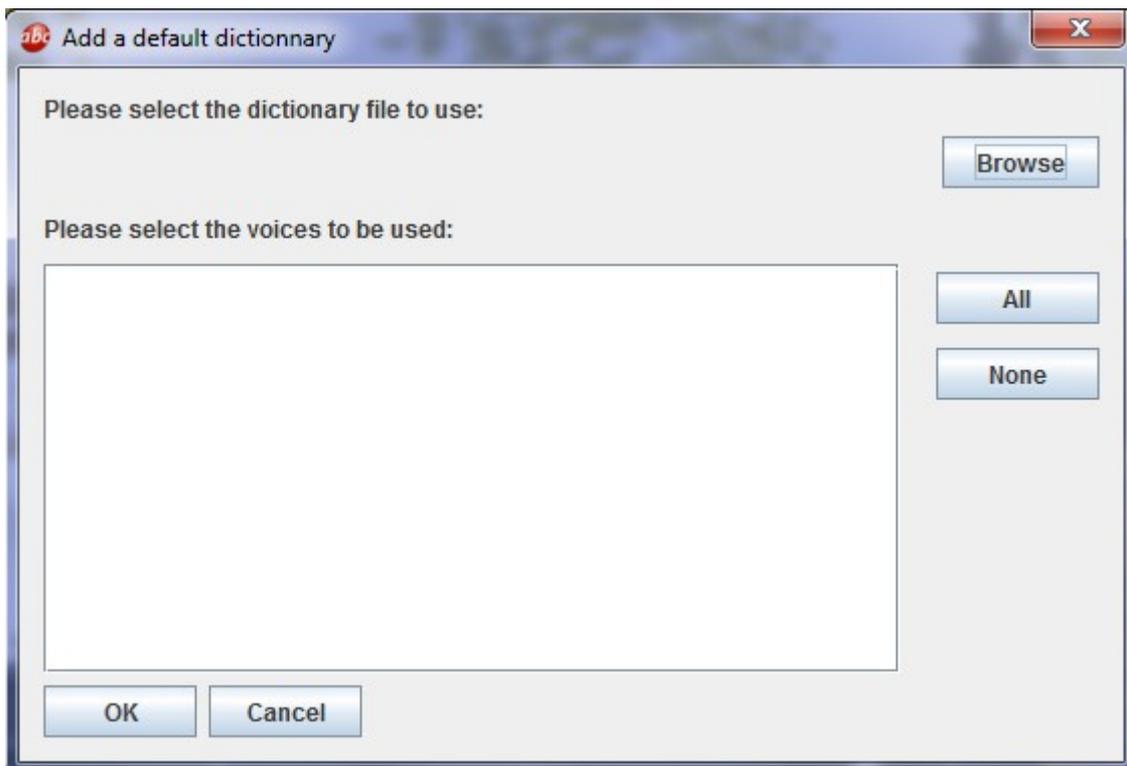
#### 6.4.5. Default Lexicons

The lexicon editor allows you to specify lexicons to be loaded and used automatically when you select a voice. You can change this setting via the **File → Manage default dictionaries** menu of the main window. This menu pops up the "default dictionaries" window.



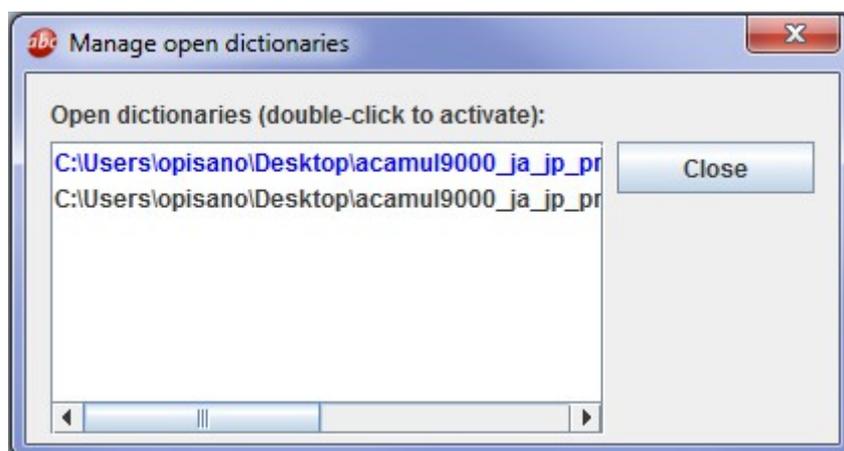
When selecting a language from the list, the window contents is updated, displaying each voice of the selected language as a directory. Under each voice node are displayed the lexicons that are loaded automatically with the voice, ordered by priority. On the right side of the window are command buttons to add, remove, and change priority order of lexicons. The **Remove**, **Move Up** and **Move Down** buttons are enable contextually, when a lexicon is selected in the list.

### 6.4.5.1. Adding a Lexicon to The Default List



This window allows you to select a lexicon to add, by clicking on the browse button and choosing a dic file from the file system. Selected file language is automatically detected and voices available for this language are displayed in the list. By clicking the **OK** button, the lexicon file is added to the default lexicon list of the selected voices.

### 6.4.6. Managing Open Lexicons



When working with several lexicons at the same time, you can end up having trouble understanding which are the ones currently loaded. The **File → Manage open dictionaries** menu is here to help you with this. It displays a window listing all the currently loaded dictionaries. The active one (the one you are currently editing) being highlighted in blue. By double-clicking one item in the list, you can change the active lexicon. The **Close** button closes a selected lexicon.

## 6.4.7. Importing/Exporting Lexicons

---

It is sometimes vital to interoperate with legacy or custom software solutions. For this, the lexicon editor provides import and export facilities with several file formats.

### Supported import file formats:

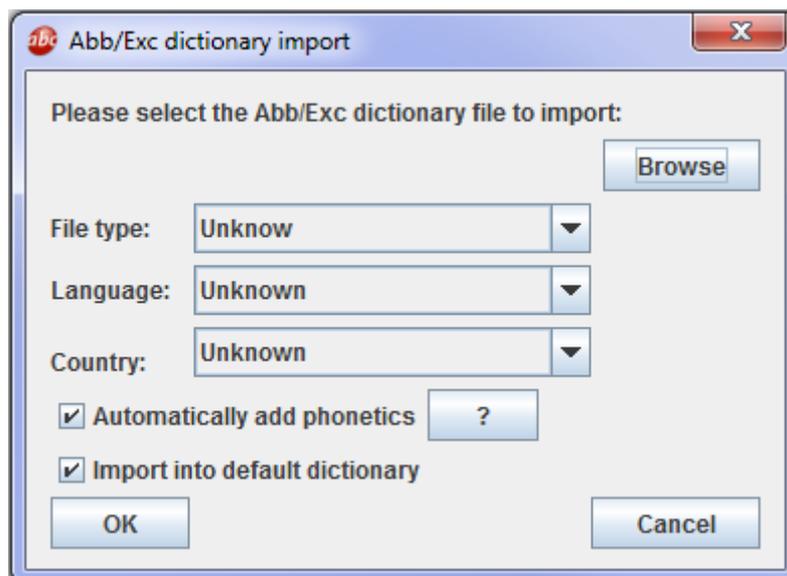
- Elan abbreviation files
- Elan exception files
- Character-separated text files
- Dictman

### Supported export file formats:

- Dictman

#### 6.4.7.1. Importing Abb/Exc Files

---

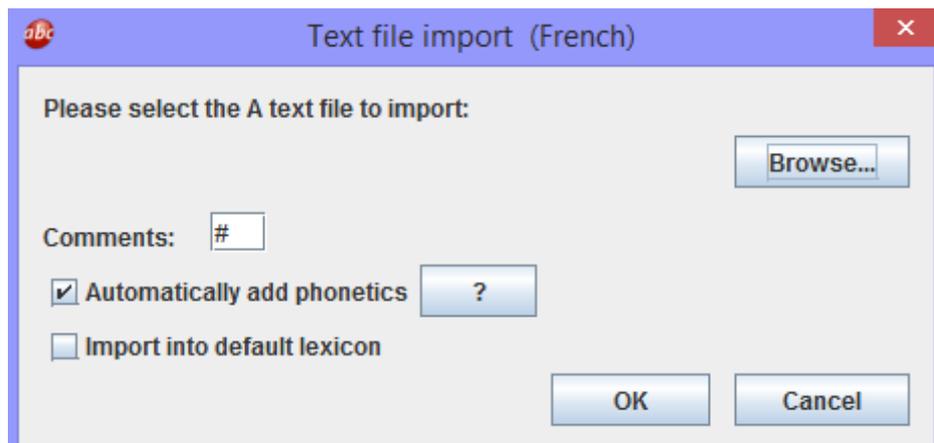


Clicking the **Browse** button enables you to select the file to import. By convention, files of this kind are named `abb_xxx.txt` and `exc_xxx.txt`. If you select a file named using the convention, the lexicon editor will try to detect the file type, language and country from the file name. Otherwise, you have to fill these informations by yourself.

The **Automatically add phonetics** checkbox plays the same role as in the Add/Modify word window: it replaces literal translations by their phonetizations.

By default, the import into default dictionary check box is checked, which means that the content of the file will be added into the language default dictionary. Otherwise you have to choose a target dic file.

## 6.4.7.2. Importing Text Files



As with the previous import window, clicking on the **Browse** button selects the file to import. Then click on **OK** to selected the target dictionary to import in (could be an existing one).

Supported text files must be composed of line-based entries with the following structure:

```
<tabulation>:wordutf8<tabulation>grammatical nature<tabulation>translationutf8
```

Where *wordutf8* is an entry and must always be prefixed by ":" (note the new dictionary text format is **utf8** encoded), *translationutf8* can be a literal or phonetic translation (the latter being prefixed by the '#' sign). The grammatical nature must be one of the followings:

Allowed natures	Meanings
NOUN	A noun
VERB	A verb
PRON	A pronoun
ADJ	An adjective
ABBR	An abbreviation
ADV	An adverb
CHIF	A cardinal
CONJ	A conjunction
FW	A contraction
DET	A determiner
INTERJ	An interjection
NUM	An ordinal
PRONIND	A quantifier
PREP	A preposition
INVARIABLE	A propernoun

---

PUNCTUATION	A punctuation
-------------	---------------

The text file can contain comment lines starting by #, which will be ignored by the import module.

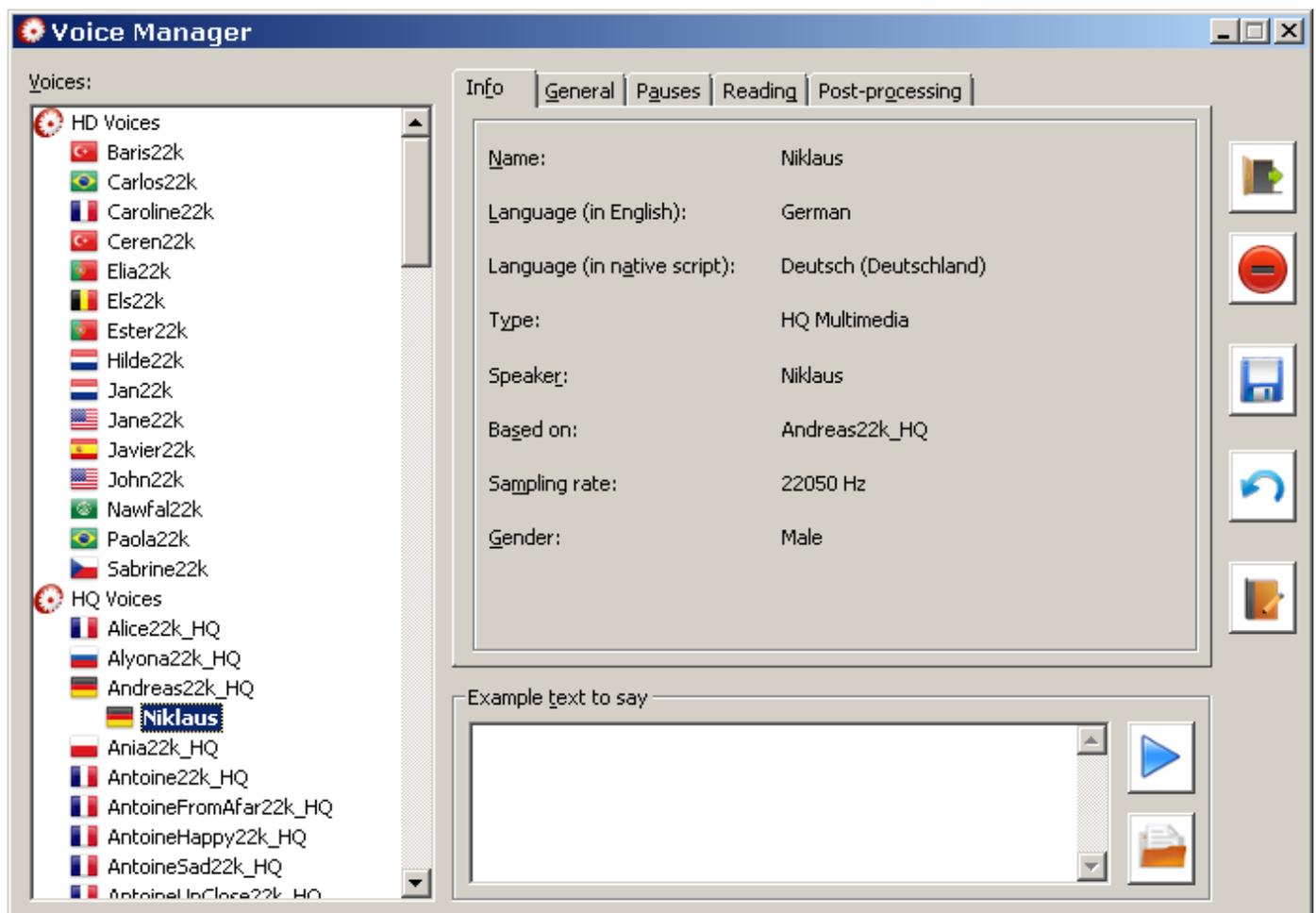
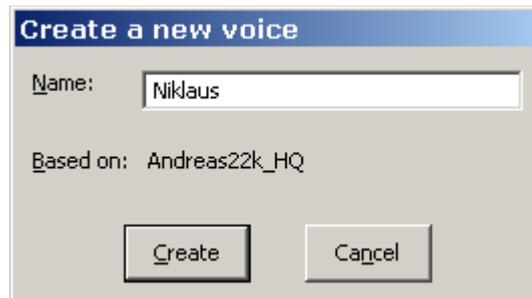
**Note:** the old text format (ref. old documentation) of dictionary can still be import in the same way, as opposite the export feature will always output text dictionary in the new utf8 format.

## 7. Voice Manager (Windows)

### 7.1. Definition

The Voice Manager is an application providing you a way to customize voices. You simply choose a base voice, change the settings of it and then save the new voice with a new name. Every AcaTTS or SAPI compliant application will then list your customized voice among the available voices.

### 7.2. The Voice Manager



#### 7.2.1. Introduction

The voice manager application (`vMan.exe`) is divided in three panels:

- **the selection tree**, used for selecting, creating and removing the voices;
- the parameters tabs, used for modifying the parameters of the selected voice (if applicable);
- **the example text box**, used for testing your changes.

### 7.2.2. The Selection Tree

---

This tree enumerates all voices you have installed on your computer:

- Voices in black are original voices.
- Voices in blue are custom voices.

With the selection tree, you may:

- Select a voice (left click).
- Create a new custom voice based on the currently selected voice (right click, context menu). This is applicable only for base voices, those that are provided as standard voices.
- Remove a custom voice (right click, context menu).

### 7.2.3. The Parameters Tabs

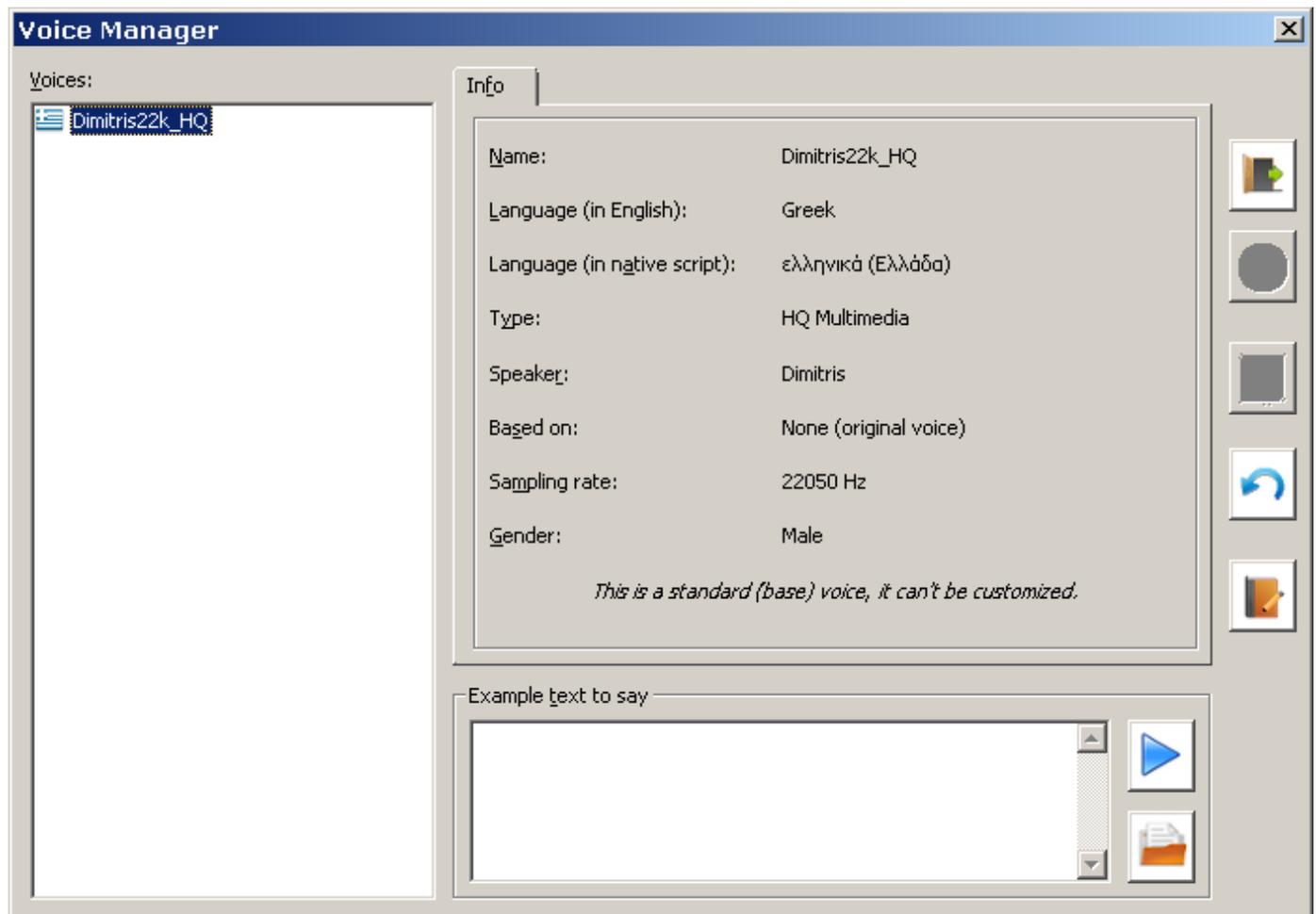
---

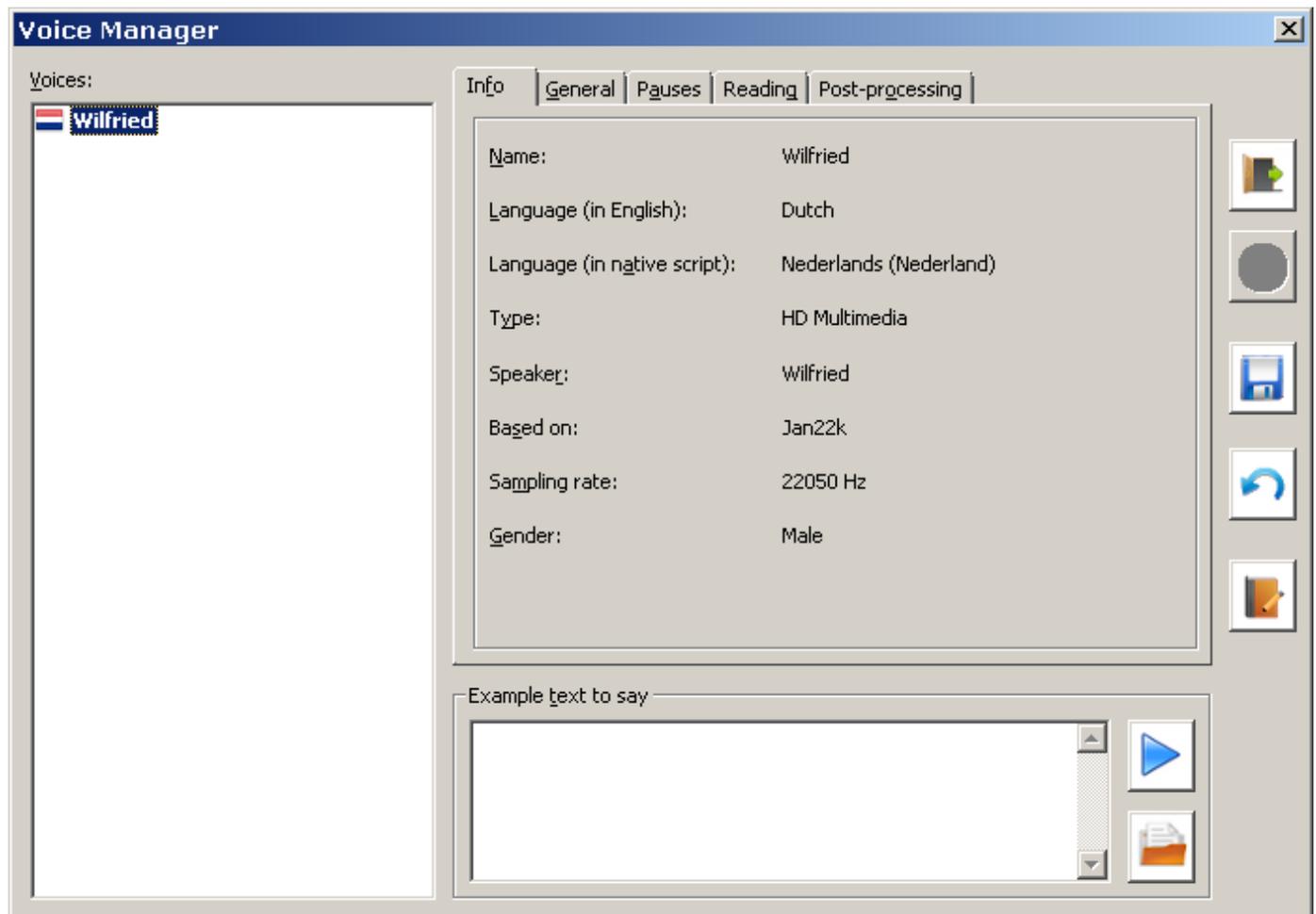
The parameters tabs are only applicable to custom voices.

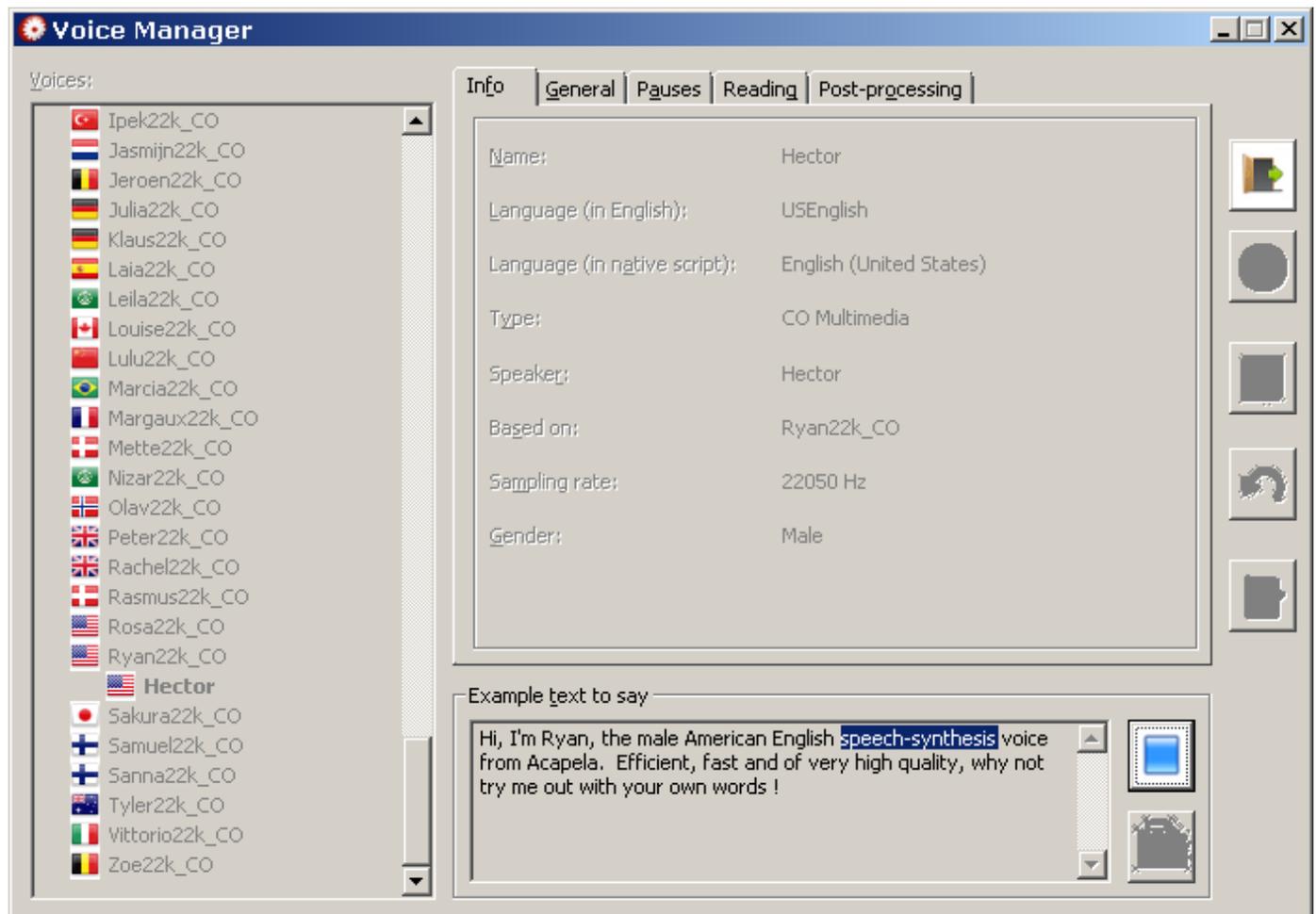
#### 7.2.3.1. Info Tab

---

This read-only tab shows you information about the selected voice.





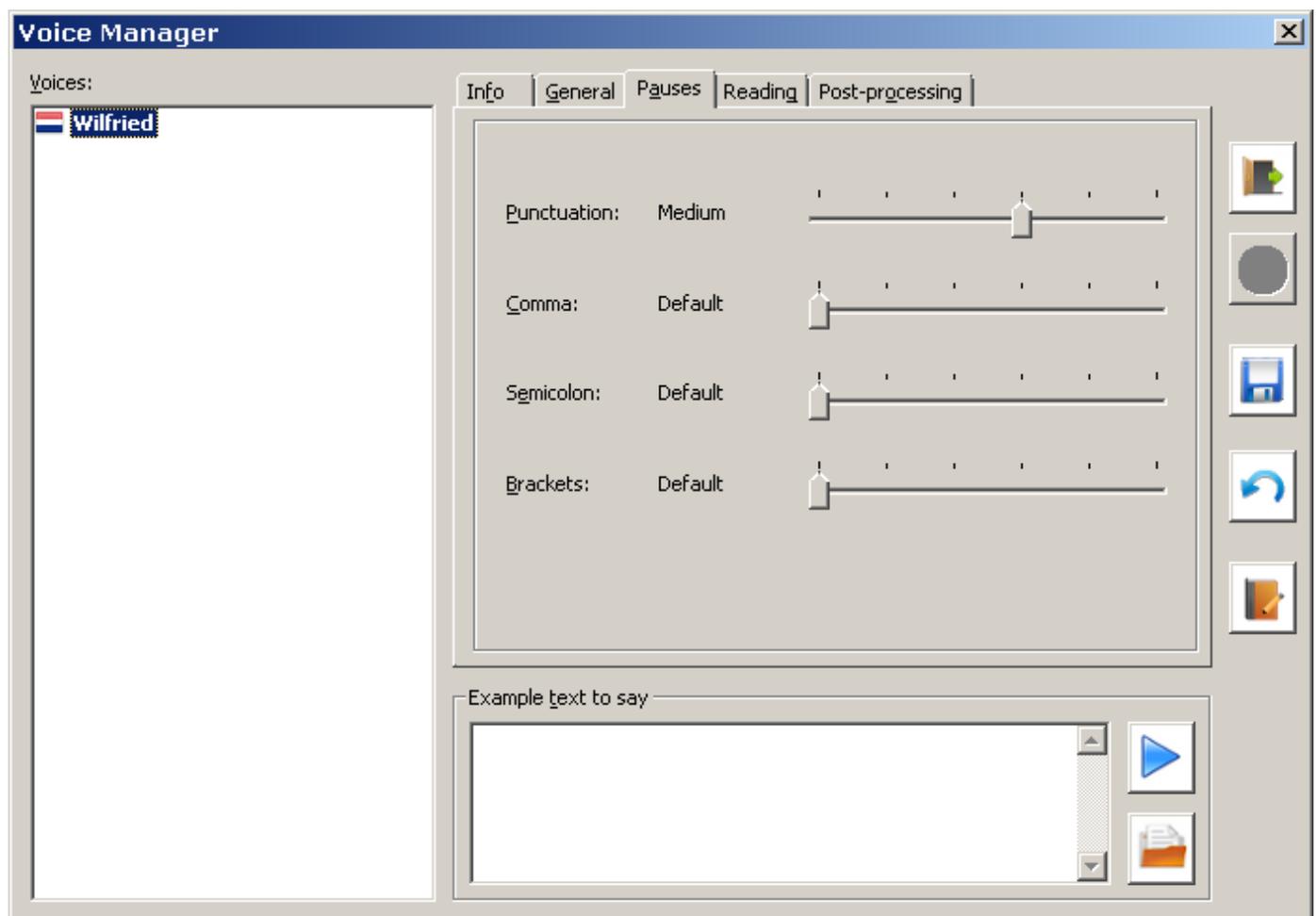


### 7.2.3.2. General Tab

Allows you to modify:

- The speed rate.
- The volume.
- The pitch (only for HD and CO voices).
- The pitch dynamic/range (only for HD voices).

### 7.2.3.3. Pauses Tab



Allows you to change the duration of each type of pause:

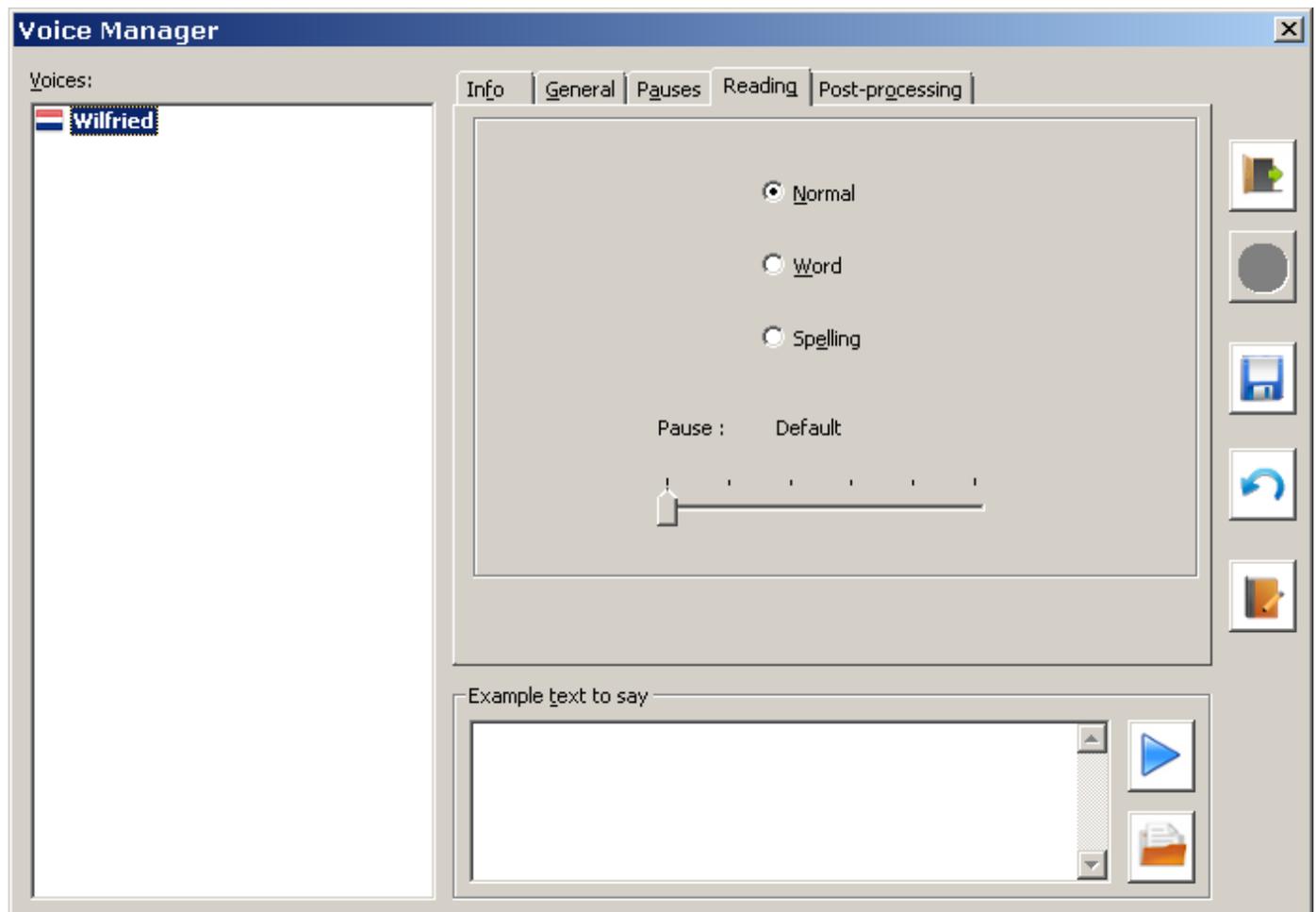
- Punctuation, end of sentence
- Comma
- Semi-colon
- Brackets

For each pause, the possible values are:

- Default: prosodic-dependent
- Very Short: 100 ms
- Short: 250 ms
- Medium: 500 ms

- Long: 1s
- Very Long: 2s

#### 7.2.3.4. Reading Mode Tab

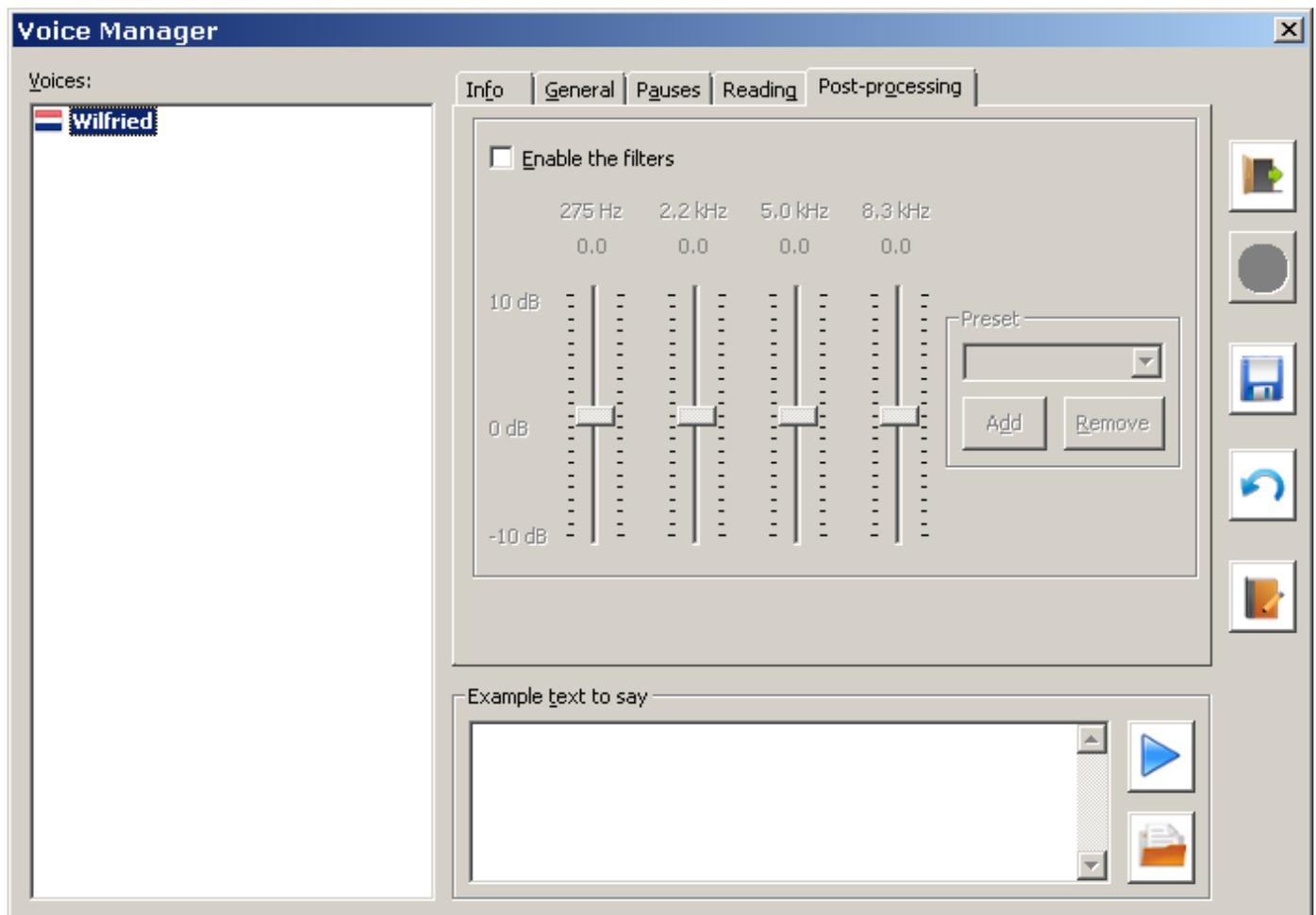


Allows modifying the way the texts are spoken:

- Normal: the text is spoken normally.
- Word: each word is followed by a pause (similar to the tag `\RmW=1\`).
- Spelling: each word is pronounced letter by letter (similar to the tag `\RmS=1\`).

For spelling and word mode, you can also modify the duration of the pause.

### 7.2.3.5. Equalizer Tab



The equalizer allows you to change the acoustical features of each voice by amplifying or attenuating each of the frequency bands.

When you start the property box the equalizer is usually off and all rulers are disabled. Click on the "Enable the filters" button to turn the equalizer on.

After you've enabled the equalizer you may simply drag the rulers to change the sound of the voice. Each ruler corresponds to a specific frequency band; the rulers on the left can be used to control the lower frequencies, while the ones on the right can be used to control the higher frequencies.

You can also use "Presets" that can be used by all voices. Simply select it with the combo box.

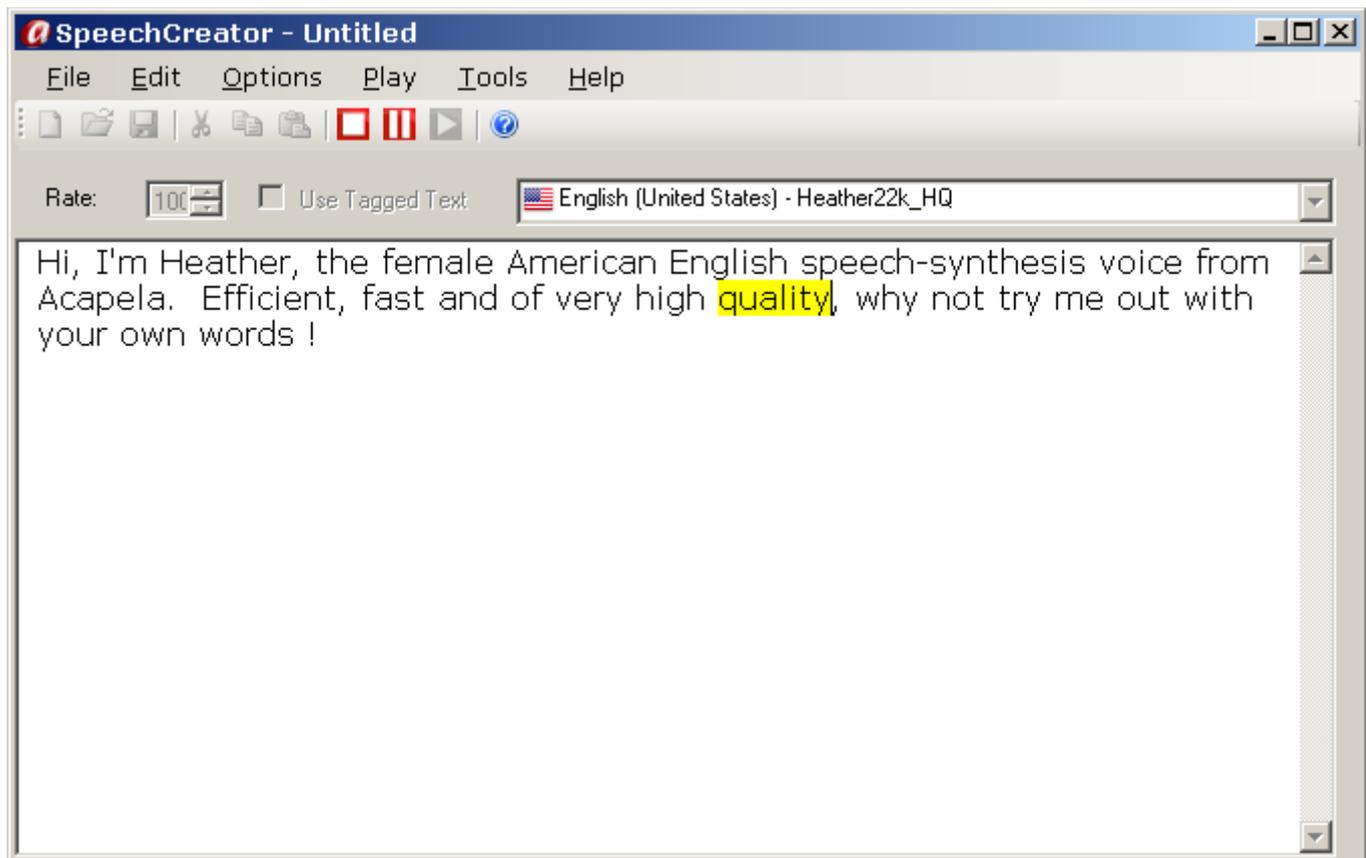
For creating a preset, you only have to adapt the rulers until you have the correct equalizer and, then typing in the combo box a new name and click the "Add" button.

You may also update an existing preset by changes the rulers and clicking on the "Update" button.

### 7.2.3.6. Other Controls

Icon	Description
	Exits the Voice Manager.
	Apply the changes permanently. Only applicable for custom voices.
	Revert : Reset the parameters to the values before change occurs.
	Open the lexicon editor for the selected voice.
	Play the text which is located in the "Example text to say" box
	Load a standard demonstration text for the selected voice (or for the base voice if the selected voice is a custom one).
	Add a custom voice based on the selected voice. Only for base voices.
	Delete the selected custom voice. Only applicable for custom voices.

## 8. SpeechCreator (Windows)



The SpeechCreator is an application demonstrating the use of Acapela TTS through the .NET binding (BabTTSNet).

It's capable of loading the demonstration texts provided with the package. It "speaks" the text and highlights the different chunks making it up (words, punctuation, etc.). It also acts as a basic editor, allowing you to copy, paste and save the text typed in the text edit zone.

Its user interface makes most of its features stand out, you can

- use tagged text; that's text that include embedded tags for inserting artificial pauses ("pau" tag), inserting a given artificial pronunciation through the specification of a string of phonemes ("prn" tag), modifying the speech rate ("spd" and "rspd" tags), etc.

See [Text Tags](#) for more information about the tags.

- modify directly the speech rate, expressed relatively to the voice's default speed. The default speech rate is 100% (neutral).
- launch the Lexicon Editor (Pronunciation Editor)
- launch the Voice Manager
- pause, resume and stop speaking the text
- select the font used to render the text

- load and save any text file, including loading and saving as Unicode files