

Getting the most out of Daisy using synthetic speech

Wolfgang Hubert
RTFC Service
Aichinger Str. 31
D-71277 Rutesheim
Germany

service@rtfc.eu
<http://www.rtfc.eu>

Abstract

RTFC is a multichannel publishing tool which has been designed to convert text documents into several accessible formats. You can produce books in common file formats like plain text, HTML and HTML Help as well as large print, braille, web braille and Daisy. It implements a standard for e-books which was created at german schools for the blind. This standard makes it possible to get the most out of Daisy even though desktop publishing software often has no capabilities to mark up optional content like annotations or sidebars. Therefore RTFC is especially suitable to convert school books and other non-fiction literature.

1 Introduction

1.1 About the history of RTFC

The history of the software is closely related to the biography of the author. The development of RTFC began in 1995. A simple program was needed to create online help from text documents. So the first version only supported the Windows Help and later the HTML Help format. This turned out to be useful for other companies as well and the software emerged from an inhouse tool to a commercial product. In 1997 the author became blind and learned braille. He decided to extend the program with large print and braille, which made it interesting for visually impaired and blind people. Finally the Daisy format has been added.

1.1 A short description of RTFC

With RTFC you can easily convert your documents into files for various media types like web pages, online help, large print, braille and digital talking books. It integrates with Microsoft Word and can be used separately to convert documents from other sources including plain text, RTF and PDF. It meets the requirements of professional publishers and teachers as well as the demands of occasional users.

There are two low-cost versions of the program: the RTFC Daisy Generator and the RTFC Braille Converter, which only support a subset of the various output formats and options. In addition you can also purchase the RTFC Personal Edition or the RTFC Professional Edition which support all output formats and features. The latter comes with additional tools to create word lists and to embed control sequences for Daisy and Braille into documents.

2 How to create Daisy books using RTFC

With the Microsoft word integration you can just click on a menu item in Word to convert a document or to launch the RTFC Hypertext Wizard. The Hypertext Wizard allows to create electronic books in a straight-forward step by step process. It offers various options for the content and layout of the target files.

2.1 Using the Hypertext Wizard

You can launch the Hypertext Wizard from the edit menu in Microsoft Word. The current document will be loaded into the Wizard automatically. You can choose the desired output format. This can be a daisy book with or without text or just MP3 files for listening on a standard MP3 player.



Illustration: Target file dialog

The fastest way to create a DTB is to choose "Format conversion" from the settings dropdown list. Click "Next" and then click "View" to start the conversion. When it is done, either an explorer window with the generated files or a browser window containing the full text will be launched depending on the document type.

2.2 Options for speech output

Before converting a document you have to choose an appropriate voice. You can choose 2 different voices which share the job of narrating your documents.

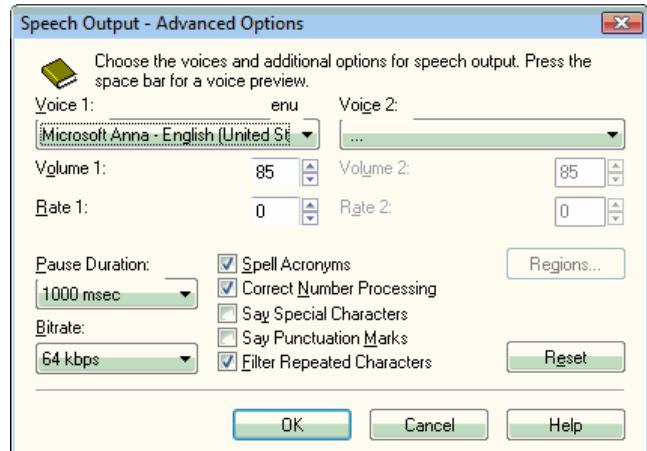


Illustration 2: Speech output dialog

You can set the volume and the speed for both voices independent from each other. The size and the quality of the audio files can be determined by the bitrate.

The drop down list for the pause duration and the various check boxes for speech options are some of the most striking features of RTFC.

With short pauses between paragraphs and other parts of a document you will get a more natural sound as from standard screen readers. The additional options allow for corrections on the native behavior of a voice. Thus you can avoid false interpretations of acronyms and prevent horizontal rules from being spoken. In addition you can correct the number processing of a voice which is especially useful to distinguish between heading numbers and dates.

2.3 Utilizing a second voice

You can choose 2 voices and determine the text regions and additional information to be spoken by the second voice.

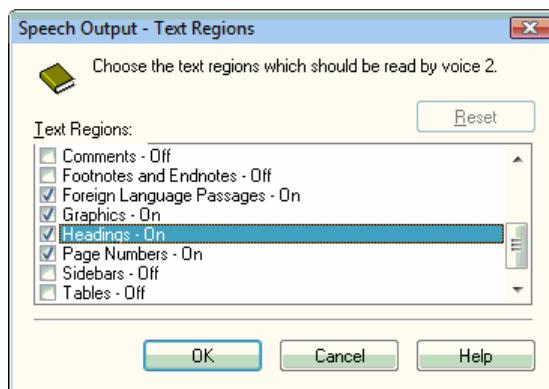


Illustration 3: Text regions dialog

This leads to the following possibilities.

1. If both voices are the same you can set a different volume and rate for the parts to be spoken by the second voice.

2. If you choose different voices for the same language (e. G. A male and a female voice in English) you can determine which parts shall be spoken by the first and the second voice. Voice 2 speaks headings and page numbers by default. This enhances the variety of your talking books. You could even let the voices change whenever there is a blank line in the text.
3. If you choose two voices which belong to a different language (e. G. An english and a german voice) you can render multiple language books. Therefore you have to prefix foreign texts by certain composition signs. This technique is well-known by braille transcribers and was adapted for the Daisy Generator.

It is also possible to let the second voice announce additional information about certain elements and the formatting of a document. This can be the following items.

- Paragraph and character styles
- The number of the row or the column in a table
- The outline lievel in a list
- The begin and end of optional content

2.4 Outline and phrases

With RTFC you can restrict the heading levels to a certain degree if there are more levels in a document than needed. You can also combine several documents to one Daisy book where the titles of the books will appear as primary headings and the topics from the documents below.

Certain parts of a document like lists, tables and graphics can be rendered as groups in a Daisy book.

Phrases can be varied from parts of sentences up to whole paragraphs.

2.5 Generating lists

Although a table of contents will not be needed in a Daisy book, RTFC generates one for the full text. So you can completely navigate a book like a web page.

You can also create an Index with links to the topics where a keyword is mentioned. In addition the page number will be announced so you can enter the number into your Daisy player and listen to the section containing the keyword.

You can even create a glossary which is an alphabetic list to all subsections in a certain section. This is not important for the audio part of a Daisy book but useful for navigating the full text.

3 Getting the most out of Daisy

There are various production tools to create Daisy books from pre-recorded material, but not as much for synthetic speech. RTFC focuses on publishing talking books without paying for human narrators. It is especially designed for converting school books and other non-fiction literature.

3.1 Document types

RTFC supports the following types of DTB (Digital Talking Book).

1. Full audio with NCC only (known as type 2).
2. Full audio and full text (known as type 4).
3. Full audio and full text with braille instead of text. It can contain either contracted or uncontracted braille.

Combining audio and braille in a Daisy book has 2 main advantages.

- If someone uses a Daisy player in conjunction with a braille display he/she can read the text in high-quality braille while listening to the audio.
- If someone wants to learn braille he/she can view the braille signs while listening to the printed text.

3.2 Handling optional content

3.2.1 The challenge: restrictions of word processors

During the development of RTFC the problem of word processors lacking functions to mark up optional content was an issue. Some word processors have facilities to move text to the sidebar of a page, but it is not apparent from the file that this is text on a sidebar. Most text processors have facilities for creating footnotes or comments, but they cannot distinguish between annotations made by the author or the producer etc.

To get the most out of Daisy, A solution was needed which could also be handled by less proficient users of text publishing software.

3.2.2 The solution: an e-book standard

During the years of 2006 to 2008 a standard for e-books was created at german schools for the blind. The purpose of this standard was to unify the methods for preparing electronic documents in order to make them interchangeable between schools. Second, it gives the blind reader of a book more information about the contents of a book and simplifies navigation. This is especially useful for mainstreaming, where blind students have to read the same books as their sighted classmates. Another benefit of this standard is a unified layout for books published in Braille or Daisy.

The standard is simple: certain parts of a document are marked by XML tags which are inserted into a document as readable text.

RTFC extends this standard to mark optional content, so that it can be rendered in accordance with the Daisy standard.

The following examples illustrate how sidebar text and annotations can be entered in a document.

```
<sidebar>This is sidebar text.</sidebar>
<annotation>This is an annotation.</annotation>
```

There is a wiki on the internet where you can learn more about the standard:
<http://www.augenbit.de/wiki/index.php?title=E-Buch>

3.3 Handling multiple voices

By default you can use 2 voices with RTFC. There might be circumstances where you would like to use more than 2 voices or where you want to change voice parameters like volume, rate and pitch at any position in the text. Therefore RTFC provides its own method for embedding control sequences into a document. For historic reasons these sequences are called "Braille Tags". In contrast to e-book tags these are entered as hidden text in order to keep the text flow for the printed document intact. Braille Tags can be entered by a tool which comes with the RTFC Professional Edition only.

Each braille tag consists of the tag itself and a name/value pair like this:

```
<braille NAME="value">
```

If a braille tag is inserted into a document, its value will be valid until the end of the document or until the same tag appears with a different value.

3.3.1 Voice parameters

The following parameters can be changed by braille tags:

- Volume in the range from 1 % to 100 %
- Rate in the range from -10 to +10 (relative to normal speed)
- Pitch in the range from -10 to +10 (relative to normal pitch)

3.3.2 Using more than 2 voices

Voices can be changed by using the "VOICE" braille tag. You can either specify the fully qualified name of the voice or a distinct part of it as follows.

```
<braille VOICE="Ryan (USEnglish) SAPI5">This is Ryan, talking US English.  
<braille VOICE="Sarah">Hier spricht Sarah, in Deutsch.  
<braille VOICE="Heather">This is Heather, talking US English.
```

With this method you can use as many voices in the same document as there are installed on your system. You could even create a radio drama this way.

3.3.3 Using multiple Languages

You can change the language independent from the voice name. This is done by using the "LANG" braille tag. You must enter an alphabetic language code as defined by iso-639. See <http://www.oasis-open.org/cover/iso639a.html> for details.

```
<braille LANG="de">Dies ist deutscher Text.  
<braille LANG="fr">C'est text français.  
<braille LANG="en">This is english text.
```

If you have installed voices from different manufacturers RTFC will choose a voice from the same manufacturer as the first voice if possible. This is done to avoid different voice characteristics in the same Daisy book.