Using Phonemes with Ivona and rSpeak Voices

NextUp.com

March 12, 2021

Overview

Versions of rSpeak voices and Ivona voices that we currently distribute at NextUp do not include support for SAPI5 phonemes. The <PRON> tag is used in SAPI5 to specify word pronunciations using phonemes, like this:

<PRON SYM = "h eh l ow"/>

Both rSpeak and Ivona voices will currently ignore this tag.

If you happen to be using our TextAloud application (https://nextup.com) there are two ways you can work around the problem, described on pages 1 and 2 of this document.

For other TTS applications, the only workaround we can document is the Lexicon Editor utility described on page 1.

Both workarounds involve using the SSML <phoneme> tag in place of the SAPI5 <pron> tag. The SSML tag does not use the phoneme symbol set used by SAPI5. It supports IPA and X-SAMPA alphabets instead.

Future versions of rSpeak / Ivona voices may include support for the SAPI5 phonemes.

ReadSpeaker Lexicon Editor Utility

Ivona and rSpeak voice installers include a Lexicon Editor utility developed by ReadSpeaker. The utility can be used to change word pronunciations, and includes support for phonemes. The lexicon editor utility supports SSML tags only, meaning SAPI5 phonemes are not supported. Either IPA or X-SAMPA can be used to create phoneme pronunciations in SSML.

You launch the ReadSpeaker Lexicon utility from the Windows Speech Control panel. On 64 bit versions of Windows you'll need to launch the correct version of the control panel, depending on whether your tts application is 32 bit or 64 bit. If you happen to be using TextAloud, TextAloud is a 32-bit application. See http://nextup.com/forum/viewtopic.php?f=15&t=6076 for instructions on launching the speech control panel.

There are two documents in your voice install directory you'll find helpful.

For Ivona voices from NextUp, the default install directory on 64 bit systems is C:\Program Files (x86)\NextUp-Ivona\, on 32 bit systems the default install directory is C:\Program Files\NextUp-Ivona\. In your NextUp-Ivona directory, locate the directory Ivona-SAPI\doc. The document ReadSpeaker-IVONA-SAPI-Lexicon Editor.pdf provides an overview of launching the lexicon editor.

For rSpeak voices from NextUp, the default install directory on 64 bit systems is C:\Program Files (x86)\NextUp-ReadSpeaker\, on 32 bit systems the default install directory is C:\Program Files\NextUp-ReadSpeaker\. In your NextUp-ReadSpeaker directory, locate the directory NextUp-rSpeak-SAPI\doc. The document ReadSpeaker-rSpeak-SAPI-Lexicons-and-Editor.pdf provides an overview of launching the lexicon editor.

IPA and X-SAMPA symbols for US English are documented in phoneset-en_us.html (Ivona) and phoneme-map-en_us.pdf (rSpeak). There is a more complete reference at http://developer.ivona.com/en/ttsresources/ssml/ssml.html. The reference at this page includes IPA and X-SAMPA phonetic alphabets for all languages supported by Ivona. There is also good information at https://www.internationalphoneticalphabet.org/

To use phonemes in the lexicon editor, you need to translate the word you want to correct to an SSML phoneme tag using either X-SAMPA or IPA alphabets. Here are examples of how you could create different pronunciations of the word “pecan” using the X-SAMPA alphabet:

<speak><phoneme alphabet="x-sampa" ph="pi:kA:n">pecan</phoneme></speak>

<speak><phoneme alphabet="x-sampa" ph="pi:"k{n">pecan</phoneme></speak>

Since X-SAMPA characters are ASCII characters they are fairly easy to use inside xml tags. However characters like the double quote character need to be written as @quot; as in the second example above. The double quote character is used to place stress on the second syllable.

You can also use IPA characters in the ReadSpeaker lexicon utility. It is more difficult because the IPA alphabet includes non-ASCII symbols. The Lexicon editor utility uses a text encoding that does not handle these IPA symbols, and they need to be encoded as hexadecimal strings.

Here are examples of pronouncing the words “try” and “test” using IPA

<speak><phoneme alphabet="ipa" ph="tɛst">test</phoneme></speak>

<speak><phoneme alphabet="ipa" ph="traɪ">try</phoneme></speak>

Note that in these examples, non-ascii IPA symbols are entered as hexadecimal characters in the tags. For example, the ɛ symbol is encoded as #x025B; for the word "test". There is a page at https://www.internationalphoneticalphabet.org/ipa-charts/ipa-symbols-with-unicode-decimal-and-hex-codes/ to help with converting ipa symbols to hex.

One thing to keep in mind is that the lexicon editor maintains separate lexicons for each voice. We don't know of any way to configure the utility to create language specific lexicons.

Using IPA Phonemes in TextAloud

If you're using the TextAloud application from NextUp.com, it is possible to use the SSML Phoneme tag to make phoneme corrections for Ivona and rSpeak voices. You can use the TextAloud pronunciation editor utility to create phoneme corrections that are voice or language specific. You must be running TextAloud version 3.0.100 or later in order for these instructions to work. You can download the latest version of TextAloud 3 from http://nextup.com/files/ta3.exe, and the current version of TextAloud 4 from https://nextup.com/files/ta4.exe. You can also visit our download page at http://nextup.com.

If you're using TextAloud 3, by default, TextAloud 3 validates that all XML markup included in text is SAPI5 markup. You will need to disable this validation. From the TextAloud menu, click Tools -> Advanced Options -> Advanced Settings. The fourth setting from the top, labeled "Perform SAPI 5 XML Validation" defaults to Yes. Toggle this setting to No, and press the OK button to close the dialog.

The above adjustment is not required with TextAloud 4.

Next, display TextAloud Pronunciation Dictionary Maintenance. From TextAloud 3, click click Tools -> Text Processing -> Pronunciation Dictionary Maintenance. From TextAloud 4, click the Dictionary button on the main toolbar, or from the menu click Control Center -> Pronunciation Dictionary Maintenance.

Click the "New Dictionary" button to create a new dictionary. Use the Dictionary Associations section at the bottom of the dialog to determine the Language, Vendor and/or Voice the dictionary will be associated with. If you have voices from multiple vendors on your system, SSML markup is not supported by all vendors. In this case, if you associate the dictionary with a Language, you should also use the Vendor dropdown to make sure the dictionary is used only by your Ivona / rSpeak voices.

After creating the dictionary you can add dictionary entries. Say you want to change the pronunciation of the word "pecan". Click the "New Entry" button to create a new dictionary entry. Set the Text Matching dropdown to "Simple Text", and use the word "pecan" (without the quotes) as the text to match on.

Set the Pronounce using dropdown to "Respell", and use X-SAMPA alphabet to create one of these these respellings:

<phoneme alphabet="x-sampa" ph="pi:kA:n"/>

<phoneme alphabet="x-sampa" ph="pi:"k{n">

You can also use the IPA alphabet, but as noted above, IPA characters are difficult to enter from the keyboard:

<phoneme alphabet="ipa" ph="pɪˈkɑːn">pecan</phoneme>

<phoneme alphabet="ipa" ph="ˈpi.kæn">pecan</phoneme>

The tags can be shortened to exclude the </phoneme> end tag if desired. The tags above can be entered like this:

<phoneme alphabet="ipa" ph="pɪˈkɑːn"/>
<phoneme alphabet="ipa" ph="ˈpi.kæn"/>

As mentioned above, use the documents phoneset-en_us.html in your Ivona install directory, or phoneme-map-en_us.pdf in your rSpeak install directrory for a list of IPA symbols supported by US English voices. The page at http://developer.ivona.com/en/ttsresources/ssml/ssml.html documents IPA and X-SAMPA symbols for all languages supported by Ivona. There is also good information at https://www.internationalphoneticalphabet.org/

Note that TextAloud uses a text encoding that handles IPA symbols, so you do not need to enter non-ascii symbols as hexadecimal characters.