phokit module

Documentation by Allan Schwade

Last Updated on 6/15/16 at 17:52:22

Contents

1   What is Phokit?

Phokit (pronounced *FOE-kit*, [ˈfoʊ.kɪt]) is a collection of methods that help automate the creation of .pho files for MBROLA, a free speech synthesizer for non-commerical projects.

2   MBROLA and the Requirements of Phokit

To generate speech, MBROLA requires a voice file and a .pho file for each utterance. Voice files can be found at `http://www.tcts.fpms.ac.be/synthesis/mbrola/mbrcopybin.html`. The .pho files are created by the user and take the following form:

```
1  p  50
2  a  200
3  t  50
4  a  50
```

Each line of a .pho file contains the segment to be generated and the duration of the segment, separated by a single space. Phokit assists users in creating large numbers of .pho files efficiently and consistently. In order to work, users must provide phokit with a spec(ification) file and a stim(ulus) list. Both are plain text documents.

2.1   the *spec file*

The spec file contains the legal set of segments that can appear in the stim list. The spec file contains at least two lines and at most four. The two required lines list the permitted consonants and vowels as key-value pairs. The keys are 'consonant' and 'vowel'. The values are the sets of legal consonants and vowels, separated by spaces. A single colon separates the key and its value. Two additional, optional lines are used to specify default consonant and vowel durations. These lines are also structured as key-value pairs separated by a colon. The keys for these lines are 'cduration' and 'vduration'. Their values are integers. Custom durations for individual segments are specified in the stim list. A sample spec file can be found below.

```
1 consonants: p t k b d g m n
2 vowels: a e i o u
3 cduration: 70
4 vduration: 125
```

2.2   the *stim list*

The stim list contains the content of all the .pho files the user wants generated. Each line of the stim list will become an individual .pho file. The segments of each line of the stim list must be separated by a single space. For example, if the user wants to generate the word 'pata', then the stim list must contain a line that reads 'p a t a'. The name of each .pho file is the segmental content of the file plus a randomly assigned number. For instance, the name of the file in the prior example might be 'pata43.pho'.

If the user wants a particular segment to have a non-default duration, the user may mark the segment with a duration tag. Duration tags take the form <dur:X>, where X is an integer specifying the duration of the tagged segment in milliseconds. Duration tags are placed directly next to the segment they modify; no space should come between the segment and the tag. A sample stim list file can be found below.

```
1 p a t a
2 m a n<dur:95> e
3 k e l o<dur:150>
4 b<dur:95> e<dur:125> n i
```

3 Class and Methods

3.1 *class* PHOKIT.phokit()

Instantiates an instance of phokit.

Attributes:
wordlist (list): Stores each line of the stim list as an item.
specs (dict): Holds the spec file key-value pairs.
durC (int): The default consonant length in milliseconds. The default is 100.
durV (int): The default vowel length in milliseconds. The default is 100.
listC (list): A list of the consonants in the spec file.
listV (list): A list of the vowels in the spec file.
counter (int): Used to assign a unique numerical value in file names.

3.1.1 loadFiles(*wordFile, specificationFile*)

Loads the stim list and spec file into a phokit instance.

Arg:
wordFile (.txt file): the stim list file
specificationFile (.txt file): the spec file

3.1.2 makeAll()

Creates a .pho file for each line in the stim list that was loaded into the phokit object.

Example:

```
1 >>> from phokit import *
2 >>> tomsvoice=phokit()
3 >>> tomsvoice.loadFiles('stimList.txt', 'spec.txt')
4 >>> tomsvoice.makeAll()
```

3.1.3 setDefault(*cduration, vduration*)

Changes the default consonant and vowel durations from 100 to the desired values.

Arg:
cduration (int): the new default consonant duration in milliseconds
vduration (int): the new default vowel duration in milliseconds

Example:

```
1 >>> print tomsvoice.durC, tomsvoice.durV
2 100, 100
3 >>> tomsvoice.setDefault(75, 80)
4 >>> print tomsvoice.durC, tomsvoice.durV
5 75, 80
```

3.1.4 make(*tagged_word*)

Makes a .pho file based on the string provided.

Arg:
tagged_word (str): the contents of the desired .pho file.

Example:

```
1 >>> tomsvoice.make('p a<dur:200> t a')
2 >>>
```

3.1.5 inspect(*tagged_word*)

Returns an easy-to-read version of a stim list line.

Arg:
tagged_word (str): the .pho file content to inspect

Example:

```
1 >>> tomsvoice.inspect('p a<dur:200> t a')
2 word:        p     a     t     a
3 dur:         −     200   −     −
```

4   In Development

- **Pitch support**. Segments in .pho files usually have a pitch specification. Marking pitch via a <hz:> tag and the option to set a default pitch in phokit instances are forthcoming.

- **Default duration specifications for each inventory segment**. Currently the user can only specify default durations for two categories, consonants and vowels. In the future, default durations for individual segments will be supported.