HQSYN16 - Task #3855

Task # 3680 (New): RA4a - Automatic error prediction

Task # 3698 (Closed): Experiment with one-class clasification for join cost enhancements

More data for artefacts collection

06.04.2016 13:38 - Tihelka Dan

Status:	Postponed	**Start date:**	06.04.2016
Priority:	Normal	**Due date:**	10.04.2016
Assignee:	Grůber Martin	**% Done:**	0%
Category:		**Estimated time:**	0.00 hour
Target version:	RA4: Automatic error prediction and signal modification		

Description

We need more data for listening tests. Especially we need to increase the coverage of rare vowels. Currently we have:

phone	total	OK	artefact
a	78	60	18
e	82	46	36
i	49	30	19
o	92	50	42
u	23	22	1
A	123	17	104
E	4	4	0
I	23	17	6
O	0	0	0
U	4	4	0

We can either try to find additional words in the corpus (shorter, though), or build "artificial" words by joining two halves of words (or words transitions) from the corpus.

History

#1 - 06.04.2016 13:41 - Tihelka Dan

- File prepare_words.py added

Adding (rather messed) script prepare_words.py which was used to select the original list of words used for the listening tests. From the full list, only words starting/ending with unvoiced consonants were used.

#2 - 27.04.2016 10:57 - Matoušek Jindřich

- Target version changed from RA1: Analysis of artifacts in synthetic speech to RA4: Automatic error prediction and signal modification

#3 - 01.08.2016 10:50 - Grůber Martin

- Status changed from New to Assigned

- Assignee changed from Grůber Martin to Tihelka Dan

I would also need a script which is used for word parts combining (and words synthesis) as it will be probably necessary to build "artificial" words.

#4 - 09.08.2016 15:17 - Tihelka Dan

- File asf2json_mix.py added

- Assignee changed from Tihelka Dan to Grůber Martin

Script asf2json_mix.py should take ASF file with individual word instances and create a set of JSON definitions which can them be passed to TTS scripting to create synthetic words for listening.

#5 - 24.11.2016 07:21 - Grůber Martin

- Status changed from Assigned to Postponed

Files

prepare_words.py	26.8 KB	06.04.2016	Tihelka Dan
asf2json_mix.py	5.06 KB	09.08.2016	Tihelka Dan