**Purpose**

We have used the recognizer created by Laurent Besacier & Company to test if this recognizer works with short audios. Our questions’ answers are very short and simple. Furthermore, potential target population are children not as Laurent’s which are adults. We want to compare default WER values with our short audios WER results.

**Default corpus’ number of words per audio**

We have analyzed number of words per each transcription at default’s corpus and those are the values (see <https://docs.google.com/spreadsheets/d/1fwMWV-Vv2YRDZxrcJlcTKTc3IjipGZ_IYRwIwbPA4TE/edit#gid=0> ). Which means that there are not any audio with less than three words and only a few of them with three words per sentence. Most of them contains 4 words per utterance or more.

It looks that maybe this recognizer is not going to work for short answers.

**Default audios origin and age**

Also at <http://www.ddl.ish-lyon.cnrs.fr/fulltext/gelas/gelas_2012_sltu.pdf> says that :

“hus, as many news websites can be found, we decided to build our own corpus based on 16 of them preselected to be strictly monolingual (avoiding

a multilingual filtering step). As in [21], we downloaded all

news pages under different format and then applied a classical

text extraction, cleaning and filtering”

All of them are **adults**

**WER values comparison**

As we saw at <https://docs.google.com/document/d/1vWTaoW7scT8BpaFENr_lMPAHAJlSF5JXGn4fG3OUl0Y/edit> , default WER values , default 7 audios WER values and our test set audios WER values are very different:

* Default test WER values are around 35%
* Default 7 audios WER values taking from the test set are around 17%, transcriptions are similar to those which the recognizer was trained by. The test corpus is similar to the train corpus.
* Our test set audios’s WER values are around 100%

This ends up with the necessity to CREATE A NEW RECOGNIZER using NEW AUDIOS created by recording CHILDREN answering to OUR TARGET QUESTIONS.

Knowing that our target users are African children with a lot of background noise speech recognition can worked unproperly.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |