

SPEX QQC REPORT

TITLE DATABASE: FARSDAT (version 2.5)
 DATABASE OWNER / PRODUCER: Research Center of Intelligent Signal Processing, Tehran, Iran
 ELRA CATALOGUE NUMBER: Unknown
 AUTHORS OF QQC REPORT: Henk van den Heuvel
 DATE: 11 April 2003
 VERSION: 1.2

SUMMARY SHEET:

Database part	Applicability (y/n)	Quality value		
		*	**	***
1. Documentation	Y			***
2. Format	Y		**	
3. Design & contents	Y		**	
4. Speech signals	Y			***
5. Annotation files	Y			***
6. Speakers	Y		**	
7. Environments	Y			***
8. Transcriptions	Y			***
9. Lexicon	Y		**	

For each applicable part a star assessment is given.

1. * This value is given if there is not a proper and reliable fit of the contents of the SLR and the information about this part as presented in the documentation.
2. ** This value is given if the documentation well accounts for the contents of the SLR. Some small deviations are permitted.
3. *** This value is given if there is no mismatch between the documentation and the contents of the SLR.

1 Quick Quality Check Report

1.1 Documentation

The most important topics should be covered and clearly described in the documentation:

Documentation is in the HELP directory in HTML-form.

- db layout and media
OK, section 1. README.TXT file
- application potential for the SLR
OK, section 3.
- directory structure and file names
OK, section 2.
- recording equipment
OK, in section 5.
- design and contents of the recordings
OK, in sections 3 and 6.
- coding and format of the speech files
OK, section 6.1
- contents and format of the label files
OK, section 6
- Speakers
Ok, section 4. There is a full score of speaker information in \INFORMATION\SPEAKER.TXT.
- recording environments distinguished
OK, section 5.

- transcription conventions
⇒ *The segmentation and labelling procedure are missing.*

- lexicon: format and transcriptions included

The directory \LEXICON contains relevant information, including a description of the phonemes used.

1.2 Format

- The file names and directory structure should correspond to the documentation
OK

1.3 Design and contents

- All mandatory items according to the documentation should be included
- Number of effectively missing files per corpus item should be appropriate

From the documentation, we infer that each speaker uttered 20 sentences, but not the same sentences. Since all speech material of a speaker per session is in one file, it is not possible in a quick check to ascertain if each speaker uttered 20 sentences.

1.4 Speech signals

- For 2 CDs of the SLR acoustic measurements on the speech files will be made, and the results reported. The acoustical measurements involved are:
 - Clipping rate
 - SNR
 - Mean amplitude

Speech files are wav file (22 KHz, 16 bit, mono, linear) and are of excellent quality.

1.5 Annotation files

- A random selection of the annotation/label files will be checked. They should be
 - Readable

- Contain the information described in the documentation

OK, they are in the PHONEME, SENTENCE, WORD directories.

1.6 Speakers

- Speaker distributions should be in agreement with documentation

The database contains speech of 99 females and 205 males.

All speakers are described in the INFORMATION\SPEAKER.TXT files.

1.7 Environments

- Environment distributions should be in agreement with documentation

The speech sounds like carefully recorded studio speech.

1.8 Transcription

- how many speech files miss an orthographic transcription?
- All non-speech markers should be described in documentation

The PHONEME directories were analysed with following results

- There is a phoneme file for each of the 6080 utterances
- Each file contains at least 11 lines

1.9 Lexicon

- The correct set of phone symbols should be used (according to documentation)
- All words in the (orth.) transcriptions should be present in the lexicon

There is a LEXICON directory that contains PDF files with the phoneme inventory used, a phoneme transcription for each word, and an orthographic (prompted) version of each sentence.

1.10 Other remarks

The database is installed by clicking a set-up file. This set-up runs smoothly. After that the program FARSDAT database can be started. I have tried to make speech and label files by using the Query option. The resulting wav-files have the indicated coding of 22 kHz, 16 bit, as mentioned in section 5.2 of the documentation.

2 Recommendations

- None