THESIS REVIEWER'S REPORT

I. IDENTIFICATION DATA

Thesis title:	**Acoustic Analysis of Psychological Stress in Speech Signal**
Author's name:	**Jana Zázvorková**
Type of thesis:	master
Faculty/Institute:	Faculty of Electrical Engineering (FEE)
Department:	Circuit Theory
Thesis reviewer:	Tomas Lustyk
Reviewer's department:	Digiteq Automotive, s.r.o.

II. EVALUATION OF INDIVIDUAL CRITERIA

Assignment	**challenging**

How demanding was the assigned project?

The assignment of the thesis is on interesting inter-disciplinary topic. There is large potential to extend experiments. It is also very important that there is a support from the Faculty of Science at Charles University.

Fulfilment of assignment	**fulfilled**

How well does the thesis fulfil the assigned task? Have the primary goals been achieved? Which assigned tasks have been incompletely covered, and which parts of the thesis are overextended? Justify your answer.

The assignment was divided into four part. The first was to conduct the research of existing literature, followed by selection of acoustic parameters that could be used to determine the level of stress. The third part analyzed recordings of phobic people and control counterparts. In the last part the results were presented and discussed.

In the part related to the theoretical introduction voiced speech is described (Chapter 2, sub-chapter 2.1). However, I am missing information about the unvoiced speech. Is that because the unvoiced speech does not play role in stress evaluation? If yes, this should be stated in the thesis and supported by references.

Methodology	**correct**

Comment on the correctness of the approach and/or the solution methods.

The nature of the experiment is quite complex and the approach to in reflects that. The student did select only acoustic parameters that were used in previous studies or verified on other speech disorders, to make this part of study simpler. With higher number of parameters, the interpretation would take longer. The parameters from DYSAN toolbox could bring new approach to already existing acoustic analysis.

Also, the conclusions of this study could help further researchers to focus on the part of the experiment which gives the most promising results (BAT task).

One of the downsides of the thesis is the number of participants (and recordings), however it should be noted that the complex nature of the experiment has an influence on it. It is probably highly difficult to collect recordings from such a complex experiment.

Positive side related to the database is, that it only consists of females. Therefore, it is balanced. And there is no additional complexity added by male participants.

Technical level	**C - good.**

Is the thesis technically sound? How well did the student employ expertise in the field of his/her field of study? Does the student explain clearly what he/she has done?

The acoustic analysis used already existing algorithms from Praat application. Also, the further analysis was performed in Matlab. So, student needed to show skill to learn how to work with new application (Praat) and some programming skills. The parameters are sufficiently described, and they could be reproduced by other researchers.

The thesis states "*Absolute values of deciding thresholds for trend monotony were set to 2 Hz, 0.5 ST,and 0.1 syll/s, every occurrence under the given threshold was evaluated as insignificant.*" However, it is not declared how these values were obtained. Why they have such a value? There is no statistical evidence for the values.

Formal and language level, scope of thesis   **C - good.**

Are formalisms and notations used properly? Is the thesis organized in a logical way? Is the thesis sufficiently extensive? Is the thesis well-presented? Is the language clear and understandable? Is the English satisfactory?

The thesis looks nice in graphical way. The text is supported by graphs and tables. As I am not English native speaker, I do not dare to evaluate the level of English.

One of the issues I see in formal part of the thesis, is that, it could be quite confusing to read through the results, tables and graphs. The experiment is multi-level (control speaker x phobic people, spider x snake, several stages of experiment [a, b, …, h], read recording before x after). In some of the experiments phobic people serve as control. It is easy to get lost. This area could be improved, make it simpler. However, it is mainly because of the nature of the experiment.

Selection of sources, citation correctness   **B - very good.**

Does the thesis make adequate reference to earlier work on the topic? Was the selection of sources adequate? Is the student's original work clearly distinguished from earlier work in the field? Do the bibliographic citations meet the standards?

The number of references in the thesis seem to be sufficient, and all of them add a piece of puzzle. Also, the thesis is clearly distinguishable from previous works.

However, there has been large number of researches performed in speech recognition area related to emotion (as one example a review by *S. Ramakrishnan - Recognition of emotion from speech: A review*). It could add more insight from different point of view.

Additional commentary and evaluation (optional)

Comment on the overall quality of the thesis, its novelty and its impact on the field, its strengths and weaknesses, the utility of the solution that is presented, the theoretical/formal level, the student's skillfulness, etc.

In one of the chapters it is mentioned "*Mean f0 [Hz] was primarily calculated to observe the effects described in previous studies as mentioned in Chapter 3.*", however, there is no such direct comparison in discussion. The current study was not confronted with previous studies. It is not clearly stated whether the results of this study correspond with trends found in previous studies or whether the assumptions were confirmed.

III. OVERALL EVALUATION, QUESTIONS FOR THE PRESENTATION AND DEFENSE OF THE THESIS, SUGGESTED GRADE

Summarize your opinion on the thesis and explain your final grading. Pose questions that should be answered during the presentation and defense of the student's work.

The thesis focuses on analysis of speech under stress in participants with phobia (snake, spider) on read and spontaneous recordings.

The thesis uses several already verified acoustic parameters, which were selected on the basis of literature research. The database of speakers in not large, but for a pivot study like this should be sufficient. And it should be noted that due to the complex nature of experiment it is difficult to collect recordings.

The conclusion of the thesis is that:

- the most promising task to reveal stress is Behavioural Approach Test (BAT).
- a trend in parameters was observed for BAT task, however statistical analysis was not performed (small dataset).

The thesis is a pilot project in this are for Czech language and it showed some promising results. However, there are also areas where improvement could be made. Hopefully, the research will continue, and it will be extended with larger dataset, other parameters and statistical analysis of results.

Questions:
1) Unvoiced speech under stress, was any research done in this area? What about non-speech events e.g. inspiration (breathing)?
2) How the thresholds for trend monotony [2 Hz, 0.5 ST, and 0.1 syll/s] were obtained? Why they were set to these exact values?

The grade that I award for the thesis is **C - good.**

Date: **6.6.2022**   Signature: