

## Review of Bachelor's Thesis

**Student:** Valenta Jakub  
**Title:** Agreements and Disagreements between Automatic and Human Speaker Recognition (id 21943)  
**Reviewer:** Matějka Pavel, Ing., Ph.D., UPGM FIT VUT

1. **Assignment complexity** **more demanding assignment**  
This bachelor thesis is more difficult, because student had to focus on 2 tasks - human verification and automated system for speaker recognition.
2. **Completeness of assignment requirements** **assignment fulfilled**  
Student fulfilled all tasks.
3. **Length of technical report** **in usual extent**
4. **Presentation level of technical report** **88 p. (B)**  
The work has a logical structure and is easy to read. Sections are written in logical order.
5. **Formal aspects of technical report** **85 p. (B)**  
Generally, the work is very good with few small exceptions:
  - in "Rozšířený abstrakt" there is a weird part of the sentence "té doby bylo mohli"
  - Section 2.1 "Marie Svobodova lecture" - here I miss a citation or footnote about the lecture - when and where it was, where are the materials ....
  - Figure 2.2 there are Czech labels on the picture
  - Section 2.2.2 "achieve a recognition accuracy of up to 95" - % is missing.
  - Section 3.1 "Both these tasks are solved, researched, ...." It is for sure not solved ... people are still working on these topics
  - Generally unify section/Section/figure/Figure/Chapter/Image - it has to be unified in terms of letters. Examples table 4.1 (page 26), Table 5.1 (page 30). There are also cases without defining what it is like "see 2.2" or "see 4.1" - I suspect it is section in most of the cases.
6. **Literature usage** **100 p. (A)**  
Excellent work with references.
7. **Implementation results** **95 p. (A)**  
Student did a very good work. He had to combine manual listening tests and also automatic test. He used open source toolkit Kaldi. As a user of Kaldi I know it is not that easy sometimes. I appreciate that after finishing the manual tests and its analysis he went back and verified his wrong decisions whether he would change it after knowing it was wrong.  
  
I would appreciate more information and comparison with work cited and done in [18] - BUT HASR'12 Experience
8. **Utilizability of results**  
Student extended already published work on a well known existing test set. For deeper and more thorough analysis he designed his own test set from the VOXCELEB dataset.
9. **Questions for defence**
  1. What do you mean - explain the sentence in 3.2.3: "It was found out that although large scale data behave beneficially in the PLDA classifier, they may be a big problem to work with i-vector extractor." Why?
  2. What is score calibration and why is it important for presenting scores to users in speaker recognition?
10. **Total assessment** **91 p. excellent (A)**  
Overall the work is very nice, it was pleasure to read, as it is well structured and experiments are well described.

In Brno 29. May 2019

.....  
signature