REVIEWER'S OPINION OF FINAL THESIS

I. IDENTIFICATION DATA

Thesis name:	**Effect of cochlear compression on predicted speech in noise perception**
Author's name:	**Liu Yuyang**
Type of thesis :	bachelor
Faculty/Institute:	Faculty of Electrical Engineering (FEE)
Department:	Department of Electrical Power Engineering
Thesis reviewer:	Ing. Jaroslav Bouse, Ph.D.
Reviewer's department:	externalist

II. EVALUATION OF INDIVIDUAL CRITERIA

Assignment **ordinarily challenging**

Evaluation of thesis difficulty of assignment.

The topic of auditory pathway modeling is complex and challenging to grasp. Student had to learn both modeling and physiology information which surpass bachelor studies. However, aside from learning the new topic, the assignment was relatively straightforward. The student used an existing auditory peripheral model, adjusted one parameter, and compared the results using a method previously provided by the thesis supervisor.

Satisfaction of assignment **fulfilled**

Assess that handed thesis meets assignment. Present points of assignment that fell short or were extended. Try to assess importance, impact or cause of each shortcoming.

Assignment was fully fulfilled

Method of conception **correct**

Assess that student has chosen correct approach or solution methods.

Chosen method is correct, based on Vencovsky,Bures ARO2023

Technical level **E - sufficient.**

Assess level of thesis specialty, use of knowledge gained by study and by expert literature, use of sources and data gained by experience.

As I mentioned in the evaluation of assignment difficulty, I consider entering the field of auditory pathway modeling to be challenging. Unfortunately, after reading the entire thesis twice, I still feel that the student has not fully grasped this topic.

The student was tasked with evaluating the role of nonlinear compression in the Dual Resonance Non-linear Filterbank (DRNL), which simulates the human cochlea, in understanding human speech in babble noise (cocktail party effect). For this, the student used a publicly available implementation of the auditory periphery, including the DRNL filterbank, adjusted the nonlinear path gain, and replicated the experiment presented in Vencovsky and Bures (ARO2023). The student's original contribution involved feeding the adjusted and non-adjusted models with speech, with and without babble noise, and calculating the cross-correlation between the outputs. The cross-correlation coefficient indicated the ability to understand the speech.

For the naive reader, the thesis lacks clarity in distinguishing the author's original work from what is taken from the literature. While it is appropriate to introduce the DRNL filter and the overall implementation, the level of detail provided by the author is excessive. For example, in section 2.1.3 on page 5, it is sufficient to mention that the implementation used differs from the cited article by Lopez and Poveda. However, explaining the source code in such detail, especially when it appears to be produced by a Large Language Model and is simply a broken stick compressor, is unnecessary. In the end, the change in that block was made only to accommodate the change of the inputs of the DRNL model from stapes velocity to stapes displacement.

Another example is the section on the DRNL filterbank implementation. The detailed explanation is unnecessary and, worse, it is a direct copy-paste from Poveda and Meddis's "A human nonlinear cochlear filterbank" (JASA 2001) without

proper citation. Additionally, the description of how the nonlinear coefficient is employed in the implementation used by the student differs from the one described in the copied text.

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

Assess correctness of usage of formal notation. Assess typographical and language arrangement of thesis.

I am not a native English speaker, so I will not evaluate the language proficiency. From a formatting perspective, there are several instances of widows and orphaned prepositions that should have been addressed during proofreading. Additionally, I would prefer that each chapter begins on a new page, but I am unsure of the current requirements of the department.

Selection of sources, citation correctness **D - satisfactory.**

Present your opinion to student's activity when obtaining and using study materials for thesis creation. Characterize selection of sources. Assess that student used all relevant sources. Verify that all used elements are correctly distinguished from own results and thoughts. Assess that citation ethics has not been breached and that all bibliographic citations are complete and in accordance with citation convention and standards.

The author chose relevant sources and cited them quite well throughout the thesis. However, the mark was reduced due to copying two pages of text without citing the original work (page 8-9 from from Poveda and Meddis's "A human nonlinear cochlear filterbank" (JASA 2001)).

Additional commentary and evaluation

Present your opinion to achieved primary goals of thesis, e.g. level of theoretical results, level and functionality of technical or software conception, publication performance, experimental dexterity etc.

Primary goal of the thesis was achieved, it's unfortunate that the source code of the work was not added as supplement.

III. OVERALL EVALUATION, QUESTIONS FOR DEFENSE, CLASSIFICATION SUGGESTION

Summarize thesis aspects that swayed your final evaluation. Please present apt questions which student should answer during defense.

Here I will copy-paste text from my evaluations above:

The student was tasked with evaluating the role of nonlinear compression in the Dual Resonance Non-linear Filterbank (DRNL), which simulates the human cochlea, in understanding human speech in babble noise (cocktail party effect). For this, the student used a publicly available implementation of the auditory periphery, including the DRNL filterbank, adjusted the nonlinear path gain, and replicated the experiment presented by Vencovsky and Bures (ARO2023). The student's original contribution involved feeding the adjusted and non-adjusted models with speech, both with and without babble noise, and calculating the cross-correlation between the outputs. The cross-correlation coefficient indicated the ability to understand the speech, and in this regard, the student fully fulfilled the assignment.

As I mentioned before, I consider entering the field of auditory pathway modeling to be challenging, and I appreciate the author for venturing into this field. However, after reading the thesis several times, I believe the author needs more time to study this topic to improve their understanding. The reduced mark is primarily due to the aforementioned lack of understanding, chaotic structure, lack of clarity in distinguishing the author's original

work from what is taken from the literature, and portion of text copied without proper citation. If the author continues in this field, I would suggest adding more theoretical introduction, particularly basic physiology, to future works to enhance readability.

To summarize, please use my mark as a reference point. If the student demonstrates his knowledge to the committee during the defense, I am open to leaving the decision to improve the grade up to the committee.

Questions:

1) Please define for the audience what is babble noise
2) On page 8, you wrote that you are using a 44.1 kHz sampling frequency. Did you check that this, combined with the model's non-linearity, is not causing unwanted aliasing?
3) From the results, it seems that cochlear compression is not solely responsible for the improvement in understandability of speech in babble noise with increasing sound pressure levels. Do you have any suggestions on what else might be responsible for this in the human auditory system (open discussion no need for precise answer)?

I evaluate handed thesis with classification grade **E - sufficient.**

Date: **7.6.2024** Signature: