

## Master Thesis Proposals 2021/2022

### Proposal 1:

#### Title: A Deep Learning Approach to Brain Tracking of Sound

#### Academic Supervisor:

**Industrial supervisors: Emina Alickovic & Martin Skoglund** (Adj. Assoc. Professors at LiU & Senior Researchers at Oticon A/S),

**Background:** Natural listening situations that require listeners to selectively attend to a talker of interest in noisy environments with multiple competing talkers are among the most challenging situations encountered by hearing impaired listeners. Such challenges become even more pronounced with increasing background noise and may partially be overcome by adequate hearing aid signal processing support. A key finding that helped the field to progress is that speech-evoked brain responses recorded with electroencephalogram (EEG) are modulated by listener's auditory attention, revealing selective brain tracking (BT) of the target talker. Hearing aid strategies were also found to support auditory attention in the hearing-impaired brain. However, BT methods proposed in the literature are linear and are thus sub-optimal, as the human brain is a complex, non-linear system, and cannot easily be modeled by linear methods.

**Project description:** We now want to work on new machine learning methods, e.g., (deep) neural networks models, to find better input-output representations for BT. Great challenges with EEG and audio are high dimensionality, low SNR, and low correlation ( $r < 0.2$ ). This knowledge will bring us one step closer to having intelligent hearing aids steered with our brains.

**Method:** The datasets will be provided by Eriksholm Research Centre (a part of the world-leading hearing aid manufacturer Oticon A/S). The dataset contains EEG data collected from 35 participants fitted with hearing aids. The participants were instructed to attend to one of two simultaneous talkers in the foreground mixed with multi-talker babble noise in the background.

**Program Duration:** 20 weeks, 30HP, with a flexible starting date.

**Contact:** [emina.alikovic@liu.se](mailto:emina.alikovic@liu.se) and/or [martin.skoglund@liu.se](mailto:martin.skoglund@liu.se)

#### Relevant Literature:

- [1] Alickovic, Emina, et al. "A tutorial on auditory attention identification methods." *Frontiers in neuroscience* 13 (2019): 153.
- [2] Lunner, Thomas, et al. "Three new outcome measures that tap into cognitive processes required for real-life communication." *Ear and Hearing* 41.Suppl 1 (2020): 39S.
- [3] Alickovic, Emina, et al. "Neural representation enhanced for speech and reduced for background noise with a hearing aid noise reduction scheme during a selective attention task." *Frontiers in neuroscience* 14 (2020): 846.
- [4] Alickovic, Emina, et al. "Effects of hearing aid noise reduction on early and late cortical representations of competing talkers in noise." *Frontiers in Neuroscience* 15 (2021).
- [5] Geirnaert, S., et al. "Electroencephalography-based auditory attention decoding: Toward neurosteered hearing devices." *IEEE Signal Processing Magazine* 38.4 (2021): 89-102.

## **Proposal 2:**

**Title: A multi-dimensional space-time-frequency representations in speech perception**

**Academic Supervisor:**

**Industrial Supervisor: Emina Alickovic & Martin Skoglund** (Adj. Assoc. Professors at LiU & Senior Researchers at Oticon A/S),

**Background:** Natural listening situations that require listeners to selectively attend to a talker of interest in noisy environments with multiple competing talkers are among the most challenging situations encountered by hearing impaired listeners. Such challenges become even more pronounced with increasing background noise and may partially be overcome by adequate hearing aid signal processing support. A key finding that helped the field to progress is that speech-evoked brain responses recorded with electroencephalogram (EEG) are modulated by listener's auditory attention, revealing selective brain tracking (BT) of the target talker. Hearing aid strategies were also found to support auditory attention in the hearing-impaired brain. However, BT methods proposed in the literature are entirely based on time, which has yielded a series of relevant findings. However, initially good space-time-frequency resolution is lost.

**Project description:** We now want to work at new space-time-frequency characterization of human multichannel EEG in order to develop new methods that can help us understand how speech is understood in noise. Great challenges with EEG and audio are high dimensionality and low SNR. This knowledge will bring us one step closer to having intelligent hearing devices that can track the listener's brain and automatically adjust its settings to improve speech understanding in noise.

**Method:** The datasets will be provided by Eriksholm Research Centre (a part of the world-leading hearing aid manufacturer Oticon A/S). The dataset contains EEG data collected from 22 participants fitted with hearing aids. The participants were instructed to attend to one of two simultaneous talkers in the foreground mixed with multi-talker babble noise in the background.

**Program Duration:** 20 weeks, 30HP, with a flexible starting date.

**Contact:** [emina.alikovic@liu.se](mailto:emina.alikovic@liu.se) and/or [martin.skoglund@liu.se](mailto:martin.skoglund@liu.se)

### **Relevant Literature:**

- [1] Alickovic, Emina, et al. "Neural representation enhanced for speech and reduced for background noise with a hearing aid noise reduction scheme during a selective attention task." *Frontiers in neuroscience* 14 (2020): 846.
- [2] Alickovic, Emina, et al. "Effects of hearing aid noise reduction on early and late cortical representations of competing talkers in noise." *Frontiers in Neuroscience* 15 (2021).
- [3] Viswanathan, Vibha, Barbara G. Shinn-Cunningham, and Michael G. Heinz. "Speech categorization reveals the role of early-stage temporal-coherence processing in auditory scene analysis." *bioRxiv* (2021).

### **Proposal 3:**

**Title: Audiovisual processing in the wild by using video-based lip reading software**

**Academic Supervisor:**

**Industrial Supervisor: Emina Alickovic & Martin Skoglund** (Adj. Assoc. Professors at LiU & Senior Researchers at Oticon A/S),

**Project Description:** This project will develop a novel machine learning based lip-reading technique capable of classifying the discrete utterances without having access the acoustic signals (i.e., speech). The aim of this project is to recognize the words coming from a talking face, given access only to the video but not to the audio files. The second aim is to check whether the proposed method generalizes across different speaker variations. Our ultimate is to use methods developed in this project to obtain audiovisual speech from the brain activity recorded with electroencephalography (EEG) instruments.

**Program Duration:** 20 weeks, 30HP, with a flexible starting date.

**Contact:**

**Relevant Literature:**

- [1] Noda, Kuniaki, et al. "Lipreading using convolutional neural network." fifteenth annual conference of the international speech communication association. 2014.
- [2] Assael, Yannis M., et al. "Lipnet: Sentence-level lipreading." arXiv preprint arXiv:1611.01599 2.4 (2016).
- [3] Chung, Joon Son, and Andrew Zisserman. "Lip reading in the wild." Asian conference on computer vision. Springer, Cham, 2016.

## Proposal 4

### Title: Conversational eye patterns and hearing impairment

#### Academic Supervisors:

**Industrial Supervisors:** Martin Skoglund (Oticon), Martha Shiell (Oticon), Sergi Rotger Griful (Oticon).

**Background:** Speech comprehension and intelligibility are severely affected by the ability to process audio-visual cues in the scene. When listening conditions are degraded by increased background noise, listeners adapt their eye gaze patterns to follow a single talker: Normal hearing (NH) listeners show longer fixations in general [1], and both normal and hearing-impaired (HI) listeners fixate more often on the lower half of the talker's face [2-3]. While these two groups have been explored separately, a direct comparison of their eye gaze behaviours remains to be done. Presumably, this emphasis on the lower half of the face reflects the listener's reliance on visual information from the lips to supplement for the degraded auditory information. While this strategy may be helpful for improving speech intelligibility, it may also potentially come at the cost of missing other meaningful visual cues in the talker's face and body. Previous research indicates that such visual cues can help a listener predict the end of a talker's turn in a conversation [4]. As such, missed cues may result in more variable timing, or the delay, of the transfer of gaze to a new talker

**Project Description:** In close collaboration with experts in human hearing, vision, and cognition the master's student(s) will explore the gaze behaviour of hearing-impaired and normal hearing listeners while following a natural conversation between two talkers with variable levels of background noise. Consistent with previous research in single talker listening, we expect both groups to show changes in fixation duration and facial areas-of-interest when background noise levels change. Furthermore, we hypothesize that this change in eye gaze patterns will result in a reduced ability to follow conversational turn-taking cues, resulting in changes to the timing of saccades between talkers. Potential differences between HI and NH listeners will be investigated.

Experiments can be done at both/either the Humanities Lab and/or at Eriksholm Research Centre (ERH), using recorded audiovisual dyadic conversations and a Tobii spectrum eye tracker.

**Location:** Lund. Some visits to ERH.

**Program Duration:** 20 weeks, 30HP, with a flexible starting date.

#### References:

- [1] Examining the Role of Eye Movements During Conversational Listening in Noise. Šabić Edin, Henning Daniel, Myüz Hunter, Morrow Audrey, Hout Michael C., MacDonald Justin A. In *Frontiers in Psychology* 2020.
- [2] Speech, movement, and gaze behaviours during dyadic conversation in noise. Hadley, Lauren V., Brimijoin, W. Owen and Whitme, William M. . In *Nature Scientific Reports* 2020.
- [3] The Effect of Varying Talker Identity and Listening Conditions on Gaze Behavior During Audiovisual Speech Perception. Buchan, Julie N., Paré, Martin, and Munhall, Kevin G. In *Brain Research* 2008.
- [4] Knowing when to respond: the role of visual information in conversational turn exchanges. Lativ, Nida & Alsius, Agnès and Munhall, K. G.. In *Attention, Perception, & Psychophysics* 2018.

**Proposal 5:**

**Title: Optimal signal processing of brain signals used for automatic control of a hearing device.**

**Academic Supervisor:**

**Industrial Supervisor:** Emina Alickovic (Adj. Assoc. Professor at LiU & Senior Researchers at Oticon A/S), & Hamish Innes-Brown (Senior Scientist at Oticon A/S),

**Background:** Auditory brainstem responses (ABRs) are electrical responses from the brain that are driven by sound input. ABRs are measured using brain signals captured with EEG from electrodes attached to the scalp. The size of the ABR response scales with input sound intensity, a property which allows ABRs to have diagnostic and clinical value. Usually, ABRs are generated in response to many brief, transient sounds such as clicks or short tones. However recent research has shown that time-resolved regression approaches can be used to generate temporal response functions (TRF's) which have similar properties to the ABR [1]. So far, TRFs have been estimated using simple linear regression models. However, it is known that the real neural system is highly non-linear. It is also known that the EEG input has a very low SNR and non-gaussian.

**Project Description:** This project will focus on comparing and selecting the best possible pre-processing of the EEG input data that will result in the most accurate TRF model.

**Program Duration:** 20 weeks, 30HP, with a flexible starting date.

**Contact:**

**Relevant Literature:**

[1] Maddox, Ross K., and Adrian K. C. Lee. "Auditory Brainstem Responses to Continuous Natural Speech in Human Listeners." *Eneuro* 5, no. 1 (2018): ENEURO.0441-17.2018. <https://doi.org/10.1523/ENEURO.0441-17.2018>.