Content

Tuning of ASR Models – Training to new Environment considering the Data Privacy Aspects

The workshop wants to discuss what is needed in other ANSPs or industry suppliers intending to transfer the HAAWAII architecture to new airports/environments, e.g. to Budapest or Charles de Gaulle.

At least voice samples from the target environment are needed. Is it possible to transfer e.g. four hours of data from Madrid environment to HAAWAII partners or does this create a lot of data privacy obstacles? Is it easier to transfer training software to Madrid or are the results from non-experts training the models so bad, that no transfer will occur?

But not only training data of voice utterances needs to be transferred. Transcription and annotations are needed.

If you intend to test the HAAWAII engine on your environment, attend the workshop and tell the HAAWAII project your needs and concerns.

Readback Error Detection Assistant. What are the real challenges? How to get data of critical situations?

Stakeholder Workshops and discussions with ATCos have clearly shown, that everybody agrees (in principle) what a readback error is, if the phraseology would have been used according to the book. Full agreement, however, is missing when discussing, which (possible) readback error use cases should be brought to the ATCo's attention. And even more questions were open, when trying to get agreements on readback error use cases, which should be communicated to the pilot.

The HAAWAII project evaluated the quantitative numbers of Readback Error Detection Rates and False Alarm Rates. It was shown, that automatic understanding of ATCo-pilot-communication is

feasible with sufficient accuracy, which make the cases with detection of readback error of interest for large community even though still a high number of false alarms is obtained. It is currently not clear, if the achieved detection rate together with the false alarm rate increases safety or decreases it.

Each day we have thousands of readback error use cases in Europe and millions of communications which follow the command-readback sequence correctly. Most of them are not interesting. The pilot provides a wrong readback or forgets it. The ATCo takes notice of it, without any reaction. But in some seldom cases, these incidents result or could have resulted from the non-reaction of the ATCo. How can we distinguish the safety relevant from the irrelevant use cases? How can we share the interesting use cases without blaming the involved people or even not considering data privacy issues? Can we just anonymize the transcripts and exchange them? Do we need to substitute frequency and waypoint information?

If you are interested in a Readback Error Detection Assistant and can report of interesting use cases, join the workshop.

Application of Speech Recognition and Understanding the ATC environment

The HAAWAII project has addressed the following ASRU applications

- Callsign Highlighting of incoming pilot calls
- Pre-filling radar labels
- Integration of CPDLC and Automatic Speech Recognition and Understanding
- Workload prediction
- Readback Error Detection Assistant

What other application did you see and want to discuss? Keyword spotting? Phraseology check? How many Descend commands are issued in a certain time period?

Support of Simulation Pilots by ASRU

Speech Recognition and Understanding has been used for a long time in education of trainees. Even automatic readback is possible without any involvement of simulation pilots by text to speech engines. The problem is that young trainees must use the standard phraseology. So ASRU can even use a grammar-based approach. If ASRU fails the trainee deviates to much from standard phraseology and should not be recognized (at least in training). .

This approach does not work with ATCos who have 20 years of experience. They sometimes deviate (for good and bad reasons) from standard phraseology. Here HAAWAII with its high recognition rates even on noisy pilot data from the operational environment can be a solution.

Who should attend (from outside HAAWAII)

- ANSPs who want to use ASRU also for support of simulations pilots with simulations with experienced ATCOs and want to reduce the costs needed for sim pilots in this context. Full replacement of the sim pilots is not the aim (at least not in the near future).

What should be discussed?

- What are the challenges? Why are you not using your standard ASR system for this application? We need more details from users as simply “recognition accuracy is not good enough”.
- How can HAAWAII support this? At the end it is know-how transfer from HAAWAII results to third party supplier.

The Ethical question: What can and should Voice Utterance be used for (and for what not)?

HAAWAII has shown that Automatic Speech Recognition and Understanding can automatically generate a lot of reports, which are even correct.

- Deviations from ICAO Phraseology
- Deviation even from Accepted Phraseology Deviations
- Number of Commands
- Callsigns/transmissions per hour
- …

But not everything what is possible should be done. So ATCO is the owner of his/her personal data. What are the benefits for him/her? How can (s)he avoid that the data is used for non-agreed purposes?

Integration of ASR and CPDLC

ASR and CPDLC are not competitors anymore. They are complementary. Speech is used as a mouse or as a touch pad. It is just another means to input something into the system, which is then e.g. sent via data link to the pilot

What should be discussed?

What is missing to make this run? Is it useful if the ATCO has the choice between different input means?

Using Speech Recognition for Workload Prediction – How do we develop a useful tool for supervisors and also for air traffic controllers?

The HAAWAII project set out to demonstrate that it is possible to use speech recognition output to better understand controller workload. This was achieved by gaining confidence in the automatic output of the speech recognition software through comparison between automatically created data and manually checked data. Following this, training data from NATS from August and September 2020 was used to build a dashboard, which provides a visualisation of the data and provides a simple interface to extract relevant data.

Analysis of this data showed that certain parameters provide a good insight and correlate with the ATCOs subjective workload rating. These parameters are:

- Average number of transmissions per minute
- Average number of callsigns on frequency per minute
- Average speech ratio per minute.

In contrast a number of other parameters seem to be less valuable when trying to quantify workload and compare the result to subjective ratings. These parameters are:

- Average word count per transmission per minute
- Average length of transmissions in seconds
- Average number of spoken words per second.

An important conclusion that can be drawn from the work carried out on HAAWAII is that speech recognition data adds value to objective workload measurement. However, further validation work and integration with other sources of information will be required to reach operational feasibility.

The results of the proof-of-concept evaluation exercises are promising in terms of being able to show changes in subjective workload using objective speech recognition data. However, further work is needed to identify cut-off values that allow the parameters, based on speech recognition output, to be used for operational decision making (e.g., when to reconfigure sectors). Additionally, work is needed to develop a human machine interface (HMI) that is fit for purpose and presents the data to the supervisory staff in a meaningful way. Larger scale validation studies will also help to identify and further refine which parameters are best suited for workload prediction. Evaluation results showed for example that speech ratio and number of transmissions per minute provide better insights than words spoken per minute.

You should attend, if you have similar challenges and want to discuss these with others and if you can give hints for further offline analysis of voice recordings by ASRU.