Readback error detection assistant to increase ATM safety and reducing ATCo's workload

Prof. Dr. Hartmut Helmke (DLR)

SESAR 2020 SHOWCASE

#SESARShowcase

#EuropeForAviation

Readback and Hearback Errors

ATCo: arctic eagle one alfa descend flight level two zero zero

Pilot: descending flight level two two zero arctic eagle one alfa

→ READBACK_ERROR

1-4 % of utterances

ATCo: ice air four charlie delta direct to dexon

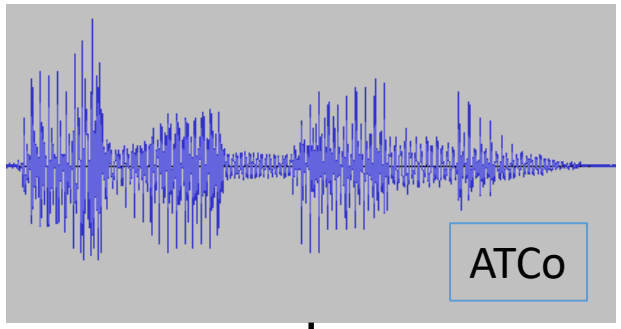
→ HEARBACK_ERROR

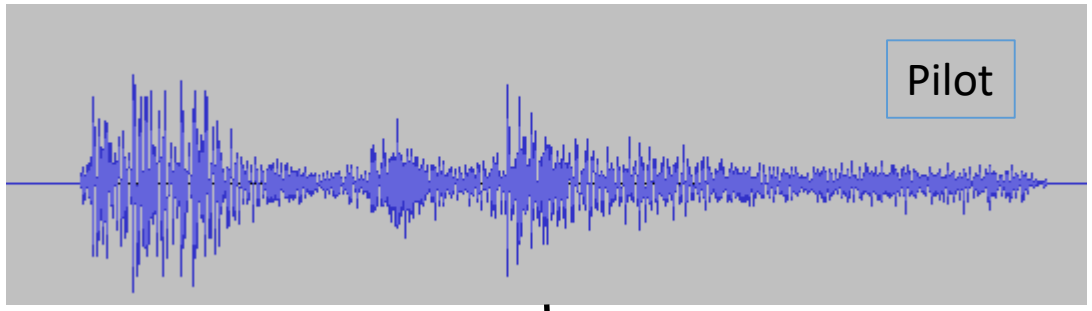
ATCo: arctic eagle one alfa negative descend flight level two zero zero

→ CORRECTED_READBACK

40-60% of readback errors result in hearback errors

Data Driven Readback Error Detection





Speech-To-Text

… flight level three six zero …

Speech-To-Text

… three four zero …

Deep Neural Network

Readback error / Readback OK

“Minor” Challenge

- We need enough training data
- 7.7 hours of transcribed and annotated voice communication resulted in 87 readback error samples

- Hard work
- Checking 2000 communication samples
- 87 samples is NOT enough

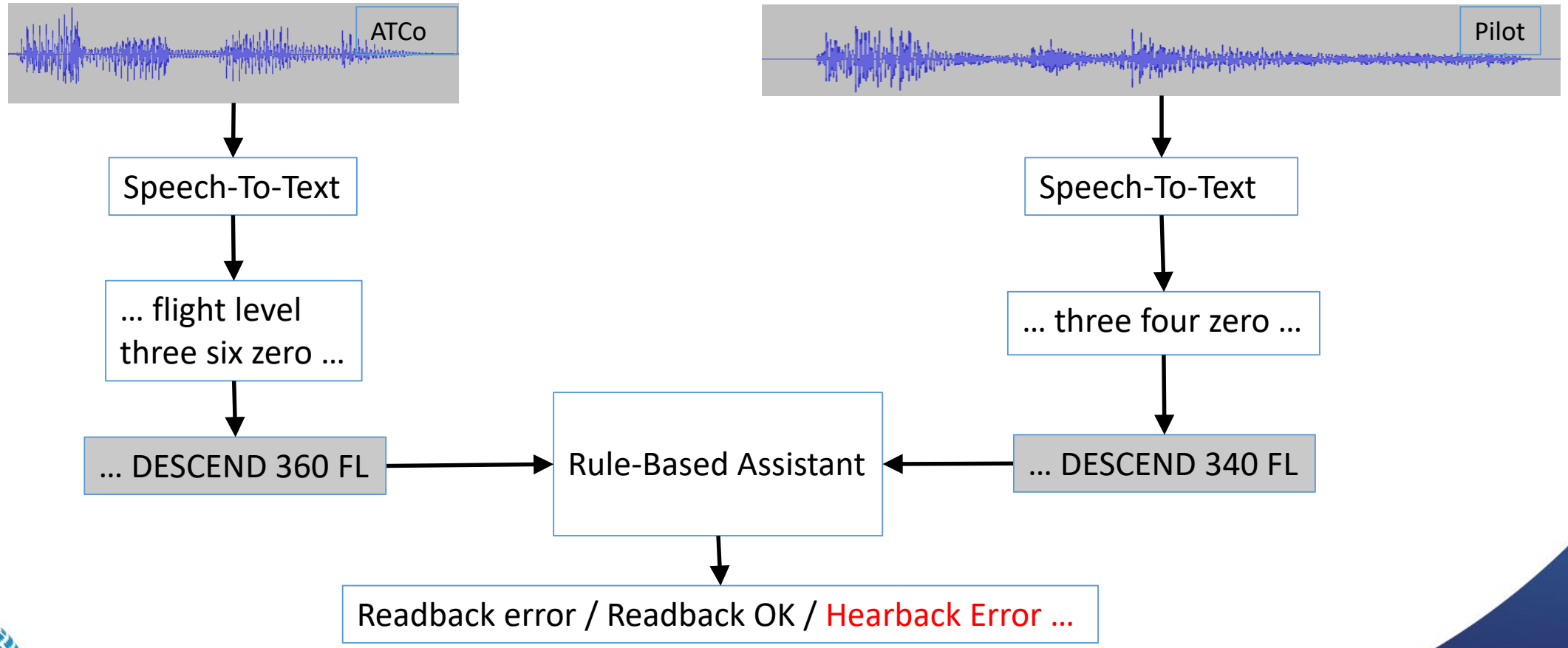
- Creating synthetic training data (data augmentation)

- 37,000 **manually** transcribed utterances from NATS and Isavia available
- 129,000 synthetic examples were created (i.e. ATCo-pilot pairs)
- 79 000 examples with readback errors for 8 different kinds
- 50 000 examples without readback errors.
- Data set split into training data (103 200 examples) and validation set (25 800 examples).

Note: All utterances used for experiments are excluded from the augmentation process.

Taken from:
H. Helmke et al., "Readback error detection by automatic speech recognition and understanding: results of HAAWAII project for Isavia's enroute airspace," in 12th SESAR Innovation Days, Budapest, Hungary, 2022.

Rule-Based Readback Error Detection



Benefits of Rule-Based Approach

FLI453 GREETING
FLI453 INIT_RESPONSE
FLI453 SQUAWK 2762

faroe line four five three good afternoon
radar contact squawk two seven six two

FLI453 SQUAWK 2761

two seven sixty one squawk
faroe line four five three

FLI453 DIRECT_TO ROBUR none WHEN PASSING GUNPA
FLI453 ALTITUDE 380 none

faroe line four five three after gunpa
cleared direct robur three eight zero

FLI453 DIRECT_TO GUNPA ROBUR none
FLI453 ALTITUDE 380 FL

gunpa robur flight level three eight
zero faroe line four five three

Data re-spoken due to data privacy issues

4:00

Results in Numbers

87 readback errors / 2200 correct readbacks

	Absolute Number	
	Detected RB Errors	False Alarms
Rule-Based	72	163

Results in Numbers

87 readback errors / 2200 correct readbacks

	Absolute Number		Rates	
	Detected RB Errors	False Alarms	Detection	False Alarm
Rule-Based	72	163	83%	69%

Results in Numbers

87 readback errors / 2200 correct readbacks

	Absolute Number		Rates			
	Detected RB Errors	False Alarms	Detection	False Alarm	Accuracy	F1-Score
Rule-Based	72	163	83%	69%	92%	45%

What is needed by the ATCo?

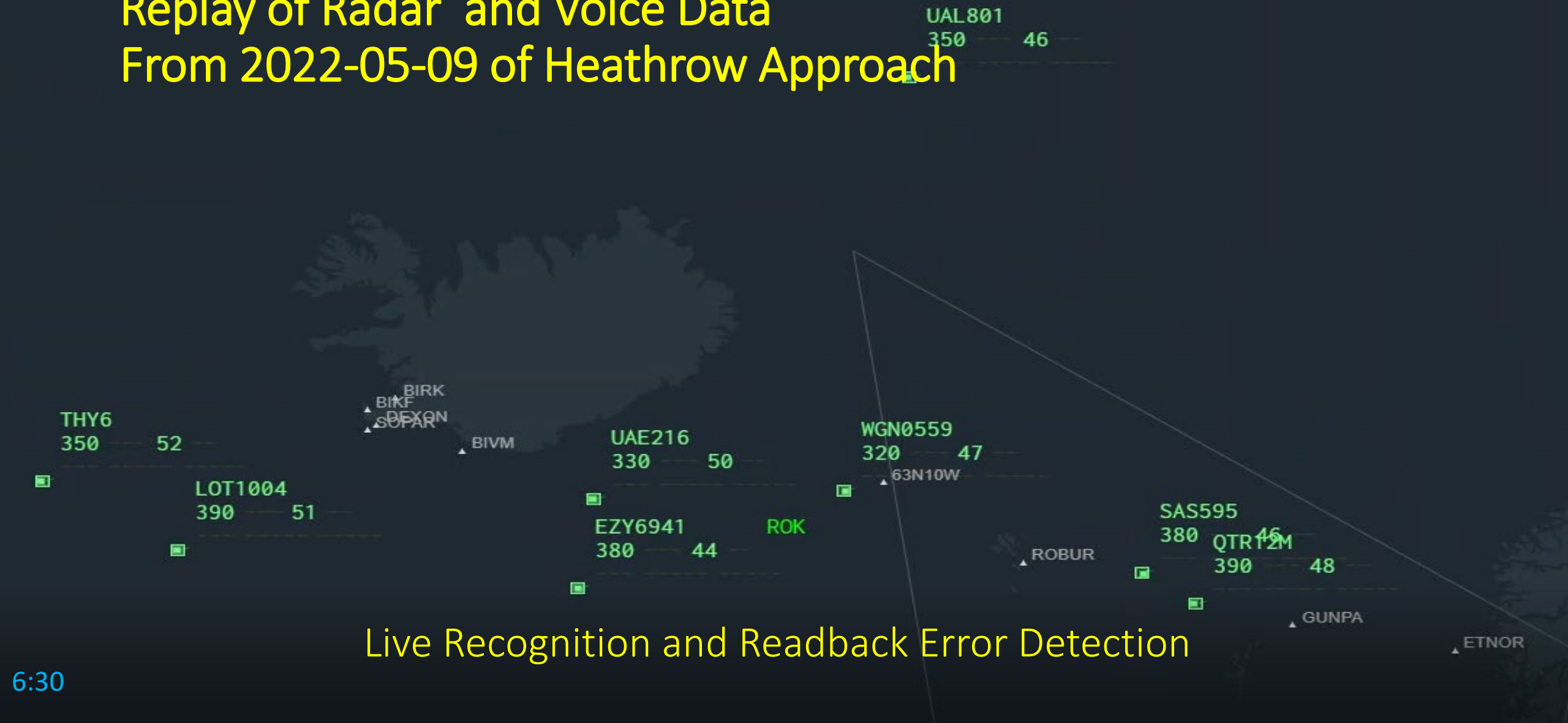
1. Readback errors **according to rules/manual**
2. Readback errors that should be **brought to the controller's attention**
3. Readback errors that should be **communicated to the pilot**

We can decide in principle. We cannot do in labs.

Readback errors are seldom cases. We need to use ops room data.

Demo at AT-One Booth

Replay of Radar and Voice Data
From 2022-05-09 of Heathrow Approach



Live Recognition and Readback Error Detection

6:30

Interactive Radar Labels

- Accept all new recognitions by click on checkmark
- Reject all new recognitions by click on cross
- After 10 seconds new recognitions are automatically accepted, if no manual or further ABSR input happens into this callsign label

User can select any output value of ABSR and manually change