



The Need for Research on Automatic Speech Recognition in Air Traffic Management

Hon. Prof. Dr. Hartmut Helmke,
German Aerospace Center, DLR
Project of AcListant®, AcListant®-Strips,
PJ.16-04-ASR-W1, MALORCA, HAAWAI



An Introductory Experiment

Mister Chairman,

Again thank you very much for the invitation and for the opportunity to present my view – my personal view – of the application of automatic speech recognition in air traffic management domain.

Let's see, if at the end of the day, I am still happy.

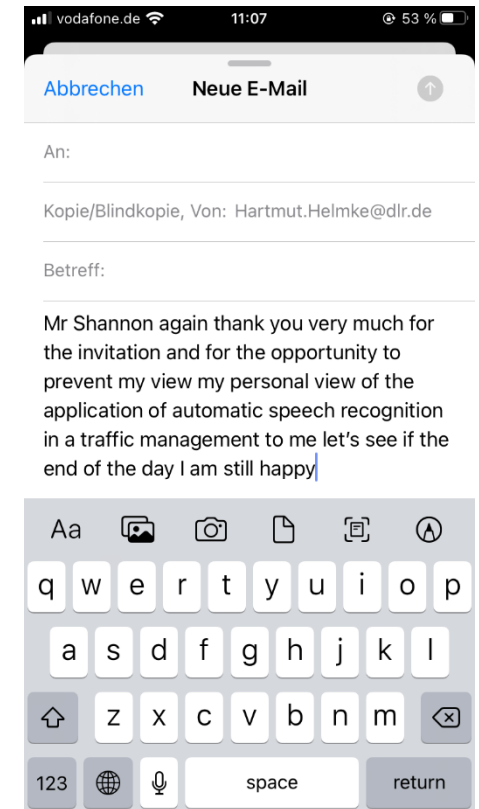


46 words

The Output of my iPhone

Mr Shannon again thank you very much for the invitation and prevent my view application of automatic speech recognition in a traffic management to me let's see if the end of the day I am still happy

Mr Shannon again thank you very much for the invitation and for the opportunity to prevent my view my personal view of the application of automatic speech recognition in a traffic management to me let's see if the end of the day I am still happy



The Output of my iPhone

HHe: Mister Chairman

4 substitutions

Iph: Mr Shannon

1 addition

HHe: again thank you very much for the invitation

1 deletion

Iph: again thank you very much for the invitation

HHe: and for the opportunity to present my view – my personal view –

Iph: and for the opportunity to prevent my view – my personal view

HHe: of the application of automatic speech recognition in air traffic management domain

Iph: of the application of automatic speech recognition in a traffic management to me

HHe: let's see if at the end of the day I am still happy

Iph: let's see if the end of the day I am still happy

Word Error Rate
 $WER = 6 / 46 = 13\%$

Why not using, Alexa, Google ... in ATM?

- Special data privacy requirements (Cloud is not an option).
- Realtime aspect are important
- ATC has a special phraseology, currently not modelled
- Special Context Information (e.g. radar or weather data) is available
- Cyber Security issues
- Safety issues, i.e. software certification for ATM application necessary
- ATM System manufacturers can't integrate ASR from Google or Amazon directly into their systems.



All pseudo arguments

Why not using, Alexa, Google ... in ATM?

lufthansa one alfa one four
descend flight level eight zero
reduce two twenty knots
contact tower one one nine decimal seven bye

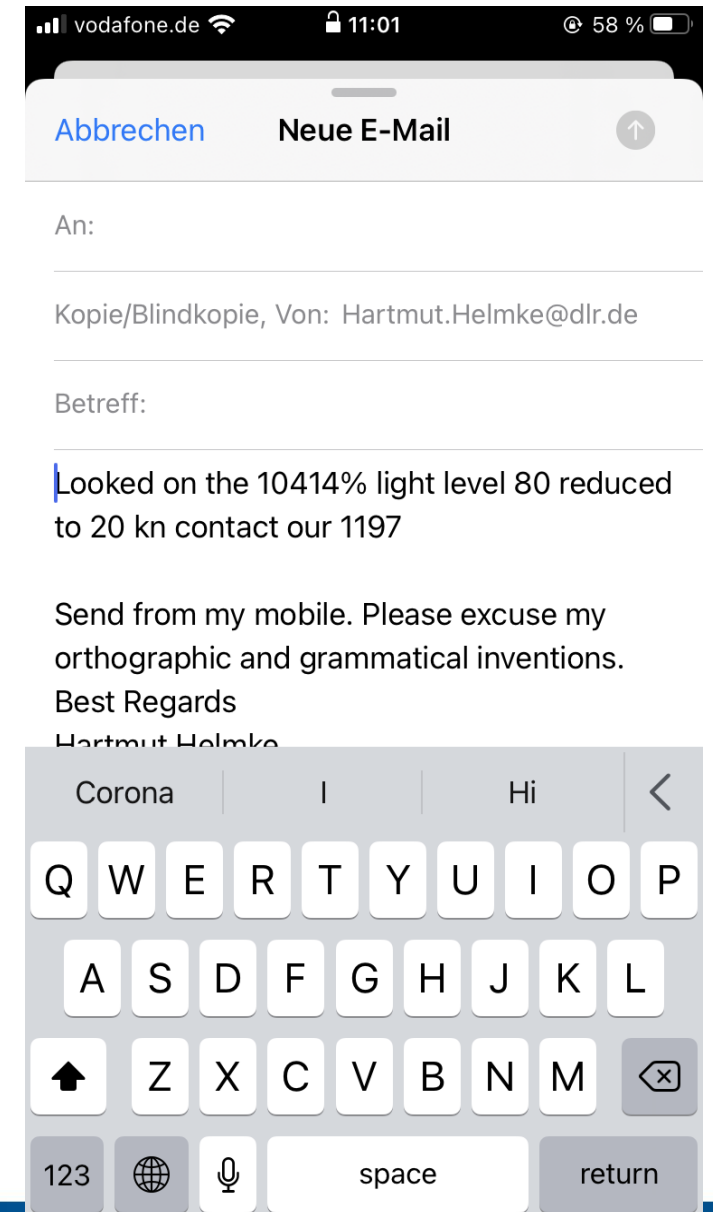


22 words



Output of my iPhone

Looked on the 10414
% light level 80
reduced to 20 kn
contact our 1197



The Output of My iPhone

HHe: lufthansa one alfa one four
lph: Looked on the 1 0 4 1 4
HHe: descend flight level eight zero
lph: % light level 8 0
HHe: reduce two twenty knots
lph: reduced to 20 kn
HHe: contact tower one one nine decimal seven bye
lph: contact our 1 1 9 7 by

9 substitutions

3 addition

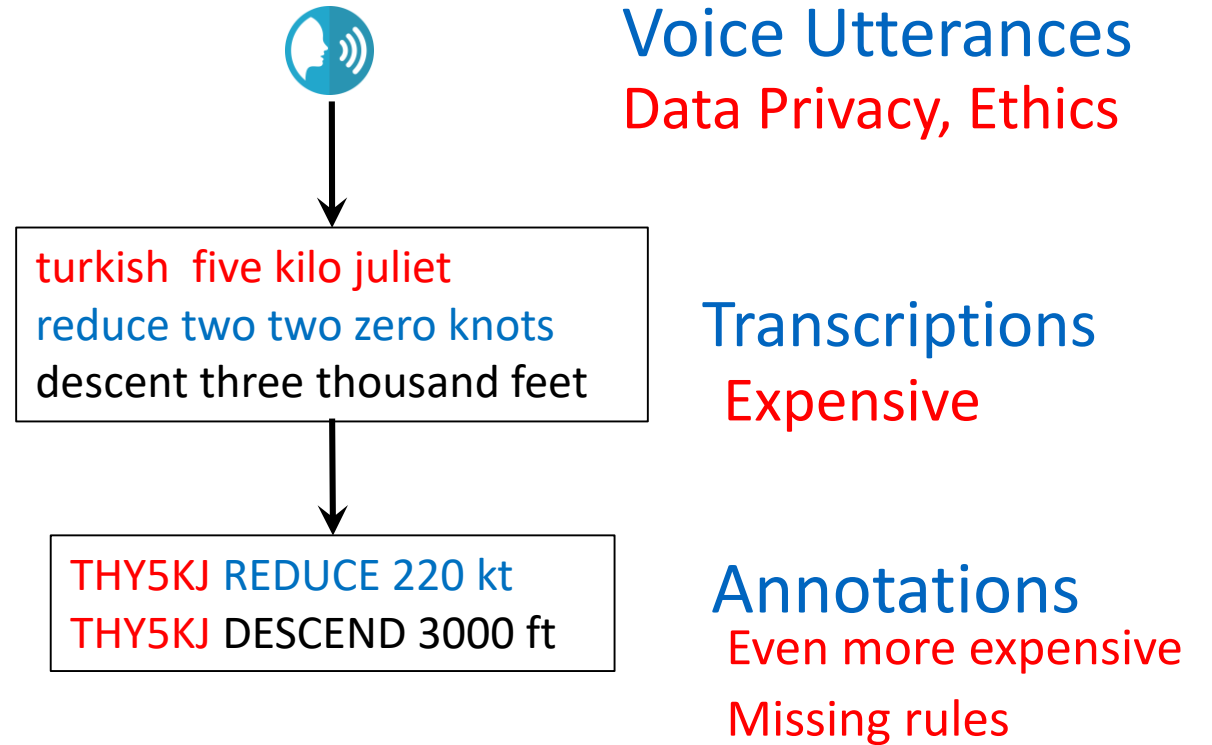
1 deletion

Word Error Rate
 $WER = 13 / 22 = 59\%$

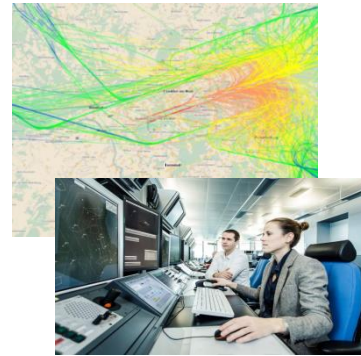
Understanding is
another challenge

What is Needed for ASR?

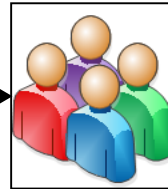
1. Most important are DATA
2. Second: DATA
3. Third: DATA



Machine Learning of Speech Recognition Models for Controller Assistance



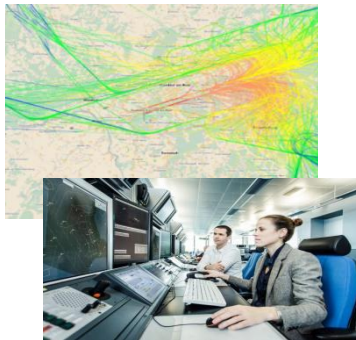
AcListant



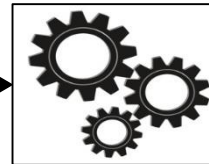
Group of experts

Automatic
Speech
Recognition

Dusseldorf



MALORCA



Automatic Learning
Process

Radar Data

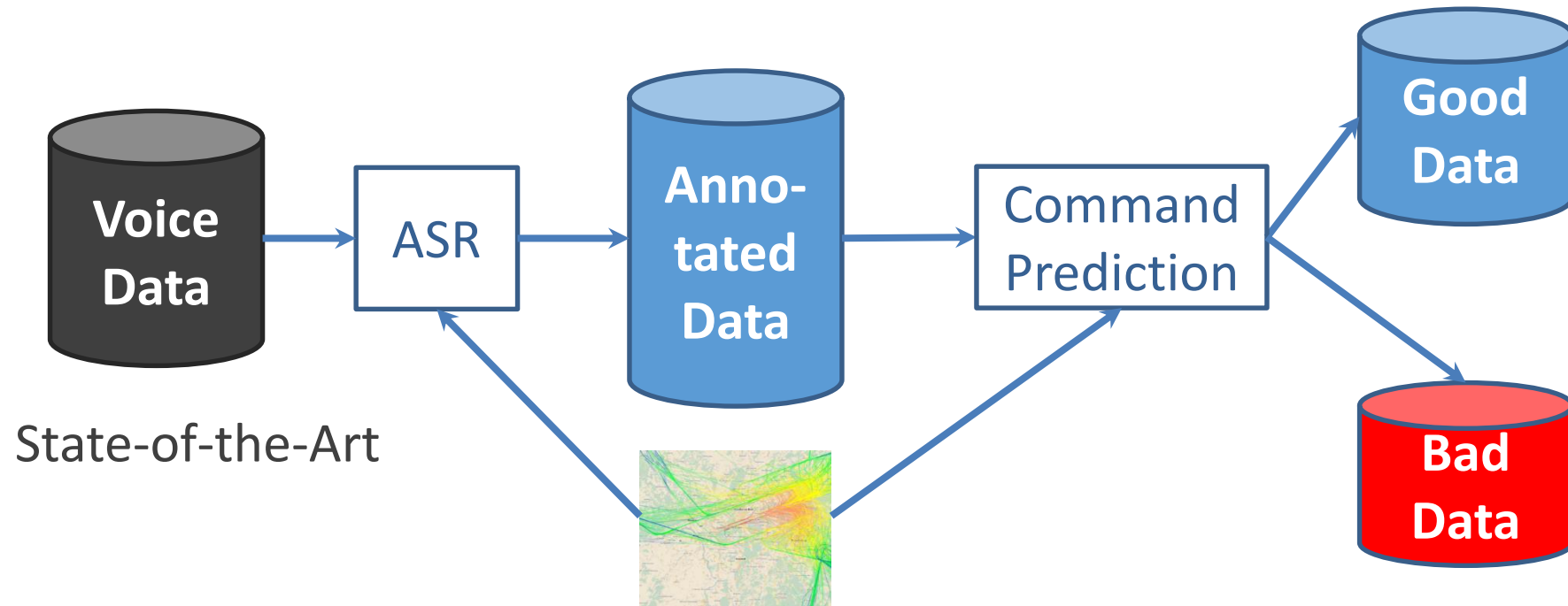
Voice Recordings

Automatic
Speech
Recognition

Prague,
Vienna

Instead of (highly skilled and paid) **experts**, **machine learning** is used.

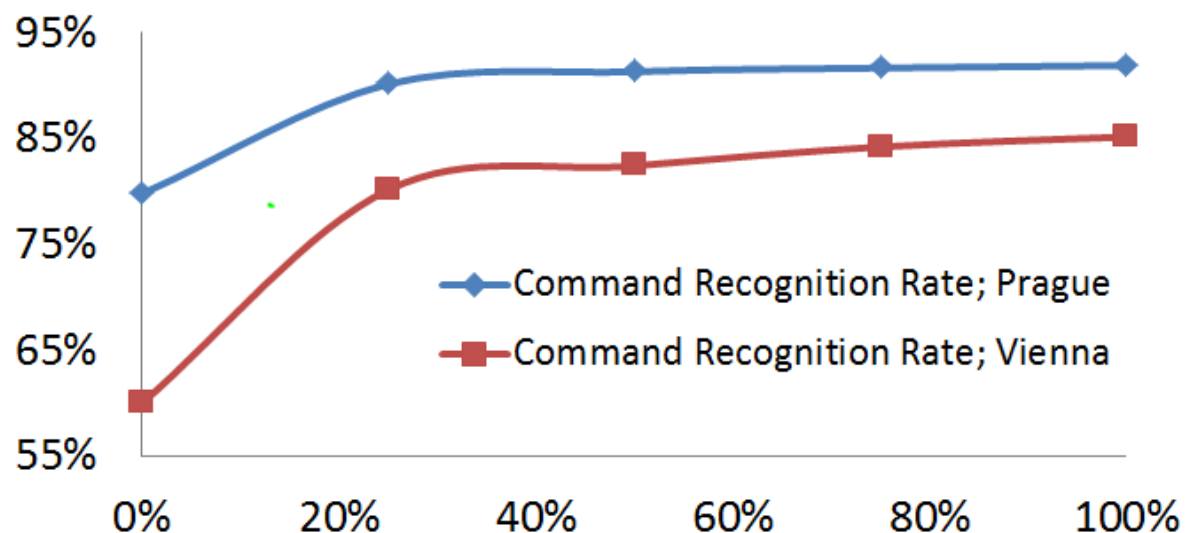
Invention of MALORCA



Data selection: Select “good” or “bad” data

Learning Curve

Command Recognition Rate (not just word recognition rate)
depending on amount of provided untranscribed training data



Vienna

Recognition Rate

0%	60.0%
25%	80.2%
50%	82.4%
75%	84.2%
100%	85.2%

WER:
5.1%

Prague

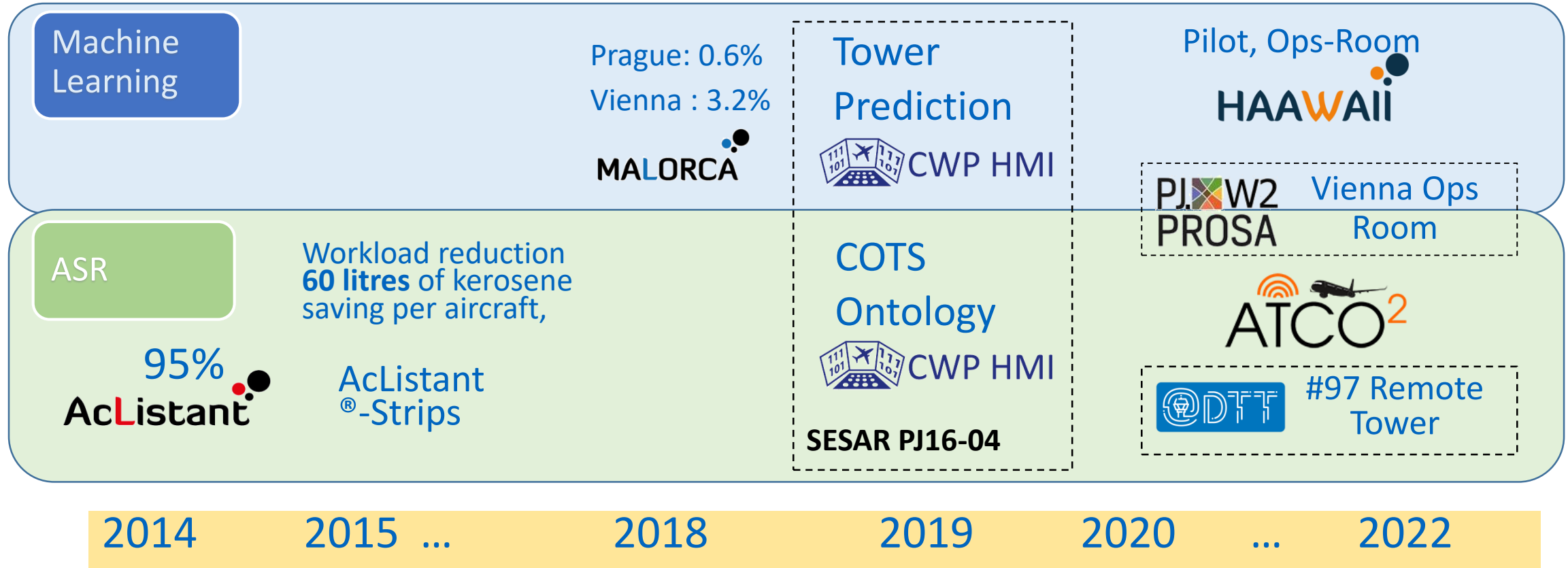
Recognition Rate

0%	79.8%
25%	90.2%
50%	91.3%
75%	91.7%
100%	91.9%

WER:
2.3%



Speech Recognition and Machine Learning Roadmap

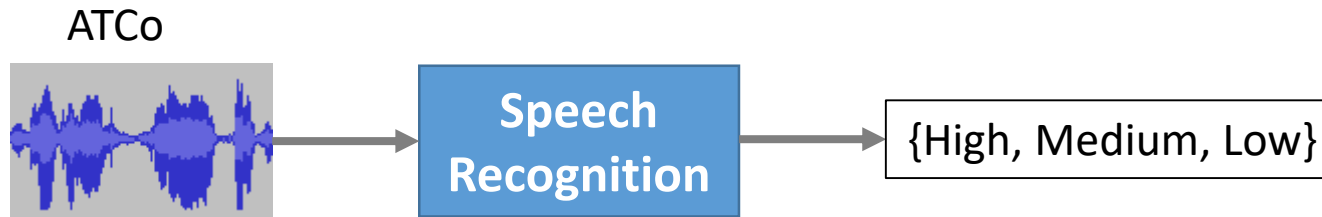


ASR Applications of HAAWAIi



Highly Automated Air Traffic Controller Workstation
with Artificial Intelligence Integration

Objective ATCo Workload Estimation



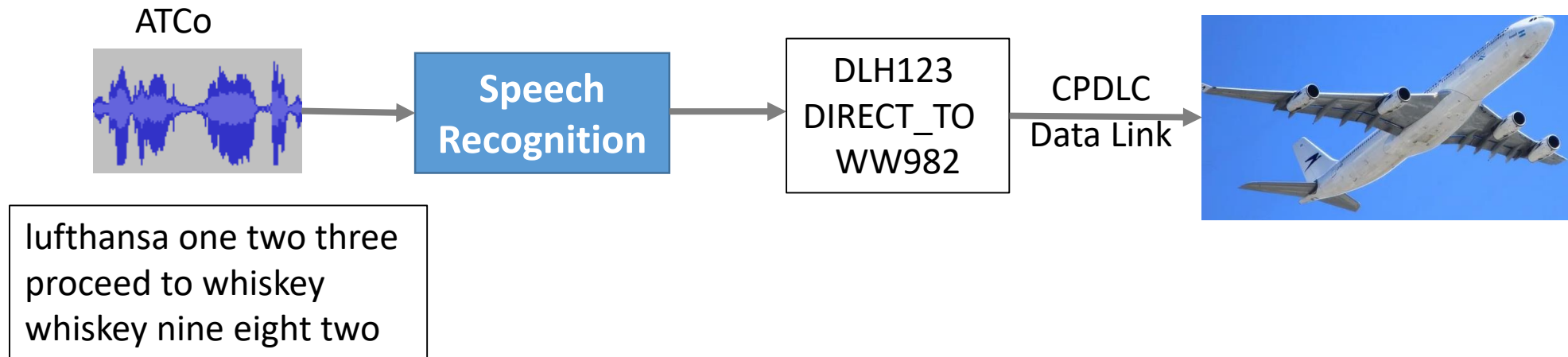
8min

ASR Applications of HAAWAIi

Integration of ASR and CPDLC



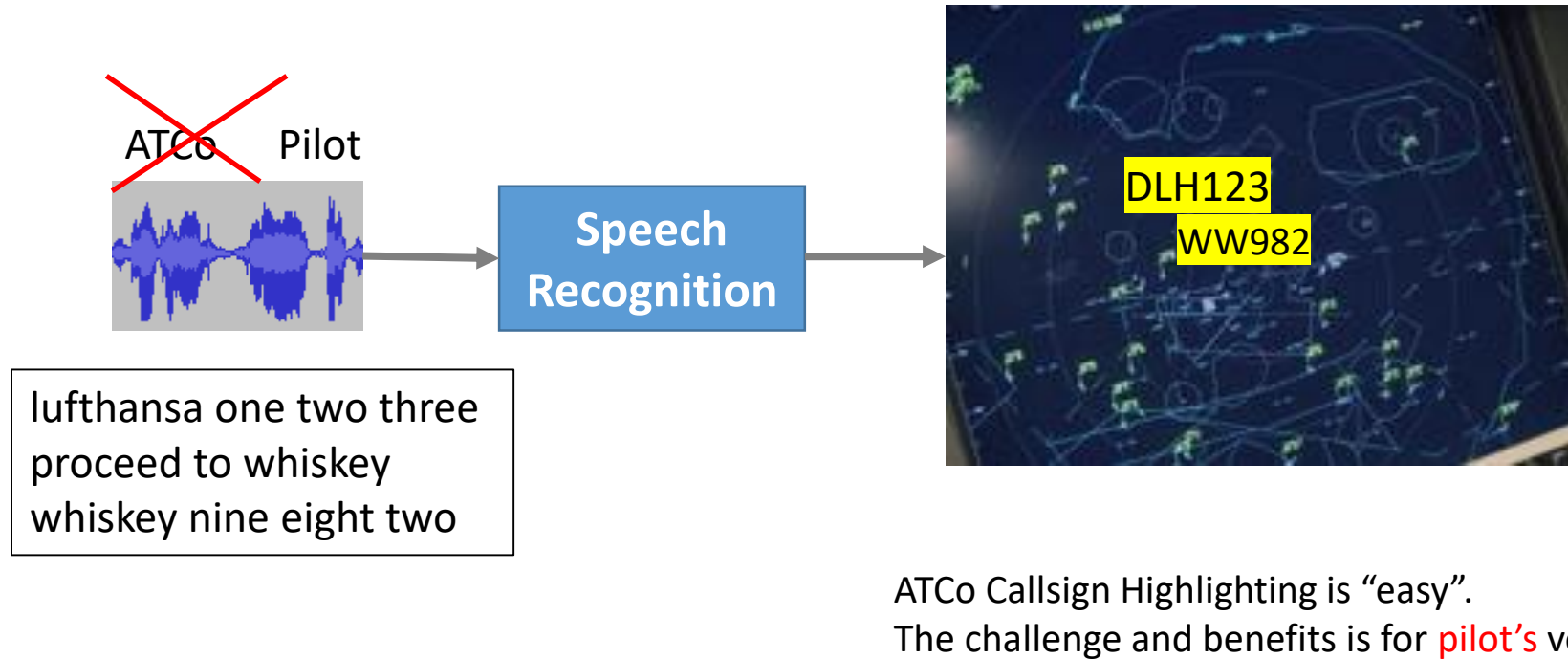
Highly Automated Air Traffic Controller Workstation
with Artificial Intelligence Integration



CPDLC: Controller –Pilot Data Link Communication

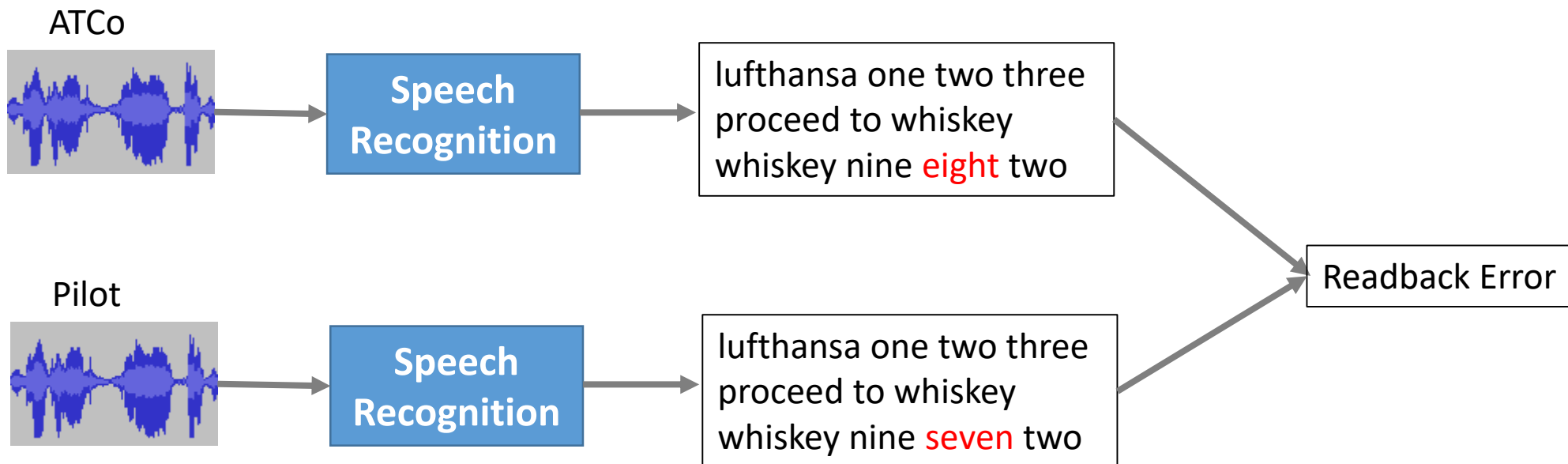
ASR Applications of HAAWAI

Callsign Highlighting & Prefilling Radar Labels



ASR Applications of HAAWAIi

Readback Error Detection (simple)



ASR Applications of HAAWAI

Readback Error Detection



Speed Recognition is NOT
Speech Understanding
Alan Turing 1952

ATCo

good morning speed bird two zero zero zero alfa
reduce one eight zero knots until DME four miles
contact tower
on frequency one one eight decimal seven zero zero

Pilot

one eighty to DME four
tower one eighteen seven
speed bird two thousand alfa

Readback Error?

- Word sequences are different
- Not each command needs a readback
- Sequence of command can be different
- “nineteen” and “one one nine” are the same
- “thousand” and “zero zero zero” are the same



ASR Applications of HAAWAI

Readback Error Detection



good morning speed bird two zero zero zero alfa
reduce one eight zero knots until DME four miles
contact tower
on frequency one one eight decimal seven zero zero

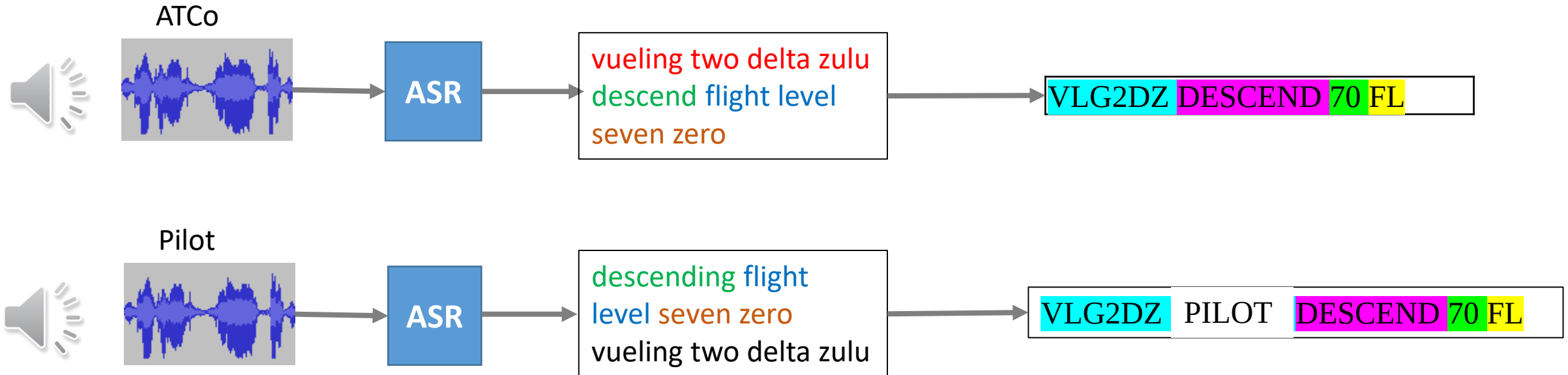
BAW2000A REDUCE 180 kt UNTIL 4 NM DME
BAW2000A CONTACT TOWER
BAW2000A CONTACT_FREQUENCY 118.700

one eighty to DME four
tower one eighteen seven
speed bird two thousand alfa

BAW2000A PILOT SPEED 180 none UNTIL 4 none DME
BAW2000A PILOT CONTACT TOWER
BAW2000A PILOT CONTACT_FREQUENCY 118.700



Understanding of ATCos and Pilots not Easy



ASR Applications of HAAWaii

Readback Error Detection

- Detection Rate > 50%
- False Alarm Rate < 10%
- 2% of Commands contain Readback errors (seldom events)

$R_{\text{both}} / E_{\text{both}}$	0.1%	0.2%	0.3%	0.4%	0.5%	0.6%
98%	4.8%	9.1%	13.0%	16.7%	20.0%	23.1%
95%	4.9%	9.4%	13.4%	17.1%	20.5%	23.6%
90%	5.2%	9.8%	14.0%	17.9%	21.4%	24.6%
85%	5.5%	10.3%	14.7%	18.7%	22.4%	25.7%
80%	5.8%	10.9%	15.5%	19.7%	23.4%	26.9%
75%	6.1%	11.6%	16.4%	20.7%	24.6%	28.2%
70%	6.5%	12.3%	17.4%	21.9%	25.9%	29.6%
60%	7.6%	14.0%	19.7%	24.6%	29.0%	32.9%
50%	8.9%	16.4%	22.7%	28.2%	32.9%	37.0%
40%	10.9%	19.7%	26.9%	32.9%	38.0%	42.4%
20%	19.7%	32.9%	42.4%	49.5%	55.1%	59.5%
10%	32.9%	49.5%	59.5%	66.2%	71.0%	74.6%

- Recognition Rates > 50% on Command Level
- Recognition Error Rates < 0.2% on Command Level

Conclusions

We need Research on Automatic Speech Recognition in Air Traffic Management

- COTS engines (iPhone, google etc.) are good, but not for ATM
- Speech Recognition does not include Speech Understanding
- Europe has an ontology
- Readback-Error detection is a challenge for research AND for subject matter experts
- Iterative approach is necessary AND possible



Thank you very much for staying in the webinar

Hartmut Helmke, Project Leader, DLR
Hartmut.Helmke@dlr.de
www.haawaii.de

