

Reducing Controller Workload by Automatic Speech Recognition Assisted Radar Label Maintenance

9th SESAR Innovation Days



Matthias Kleinert, **Hartmut Helmke**, Sylvain Moos, Petr Hlousek,
Christian Windisch, Oliver Ohneiser, Heiko Ehr, Aline Labreuil

German Aerospace Center (DLR), Thales LAS France SAS,
Air Navigation Service Provider of Czech Republic, Austro Control



Introduction

- DLR achieved command recognition rates greater 90%
- Only 75% in this exercise
 - In simulation environment
 - With Commercial-off-the-Shelf (COTS) Speech Recognizer
 - With industrial HMI
- **Lessons learned** are my message

Contents

1. Automatic Speech Recognition in SESAR context
2. Performed Exercise
3. Exercise Results
 - for Prague Approach
 - for Vienna Approach
4. Conclusions



Contents

1. Automatic Speech Recognition in SESAR context

2. Performed Exercise
3. Exercise Results
 - for Prague Approach
 - for Vienna Approach
4. Conclusions





ACG / COOPANS

ANS CR (B4)

Avinor ANS

CCL / COOPANS

CRIDA

DFS

DLR (AT-One) (Lead)

ENAIRE

FRQ (FSP)

HC (FSP)

Indra

Integra

LFV / COOPANS

LPS SR (B4)

NATS

Naviar / COOPANS

Romatsa

Thales LAS France SAS

PJ.16-04 ASR Team

ATM System Provider

ANSP

Research Institute

Consultancy



founding members





PJ.16-04

5 exercises related to Automatic Speech Recognition

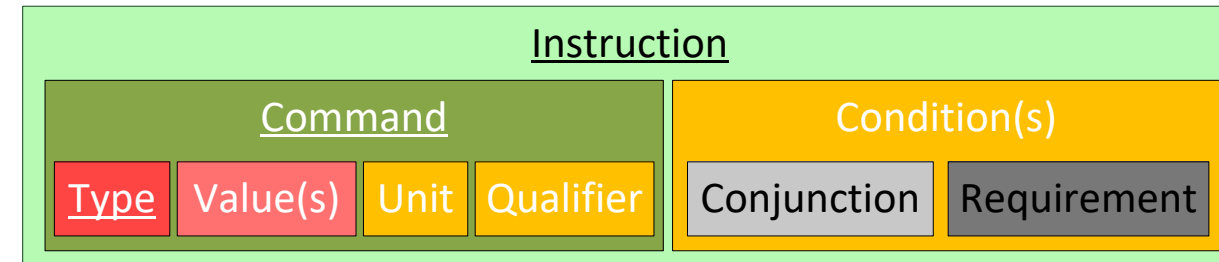
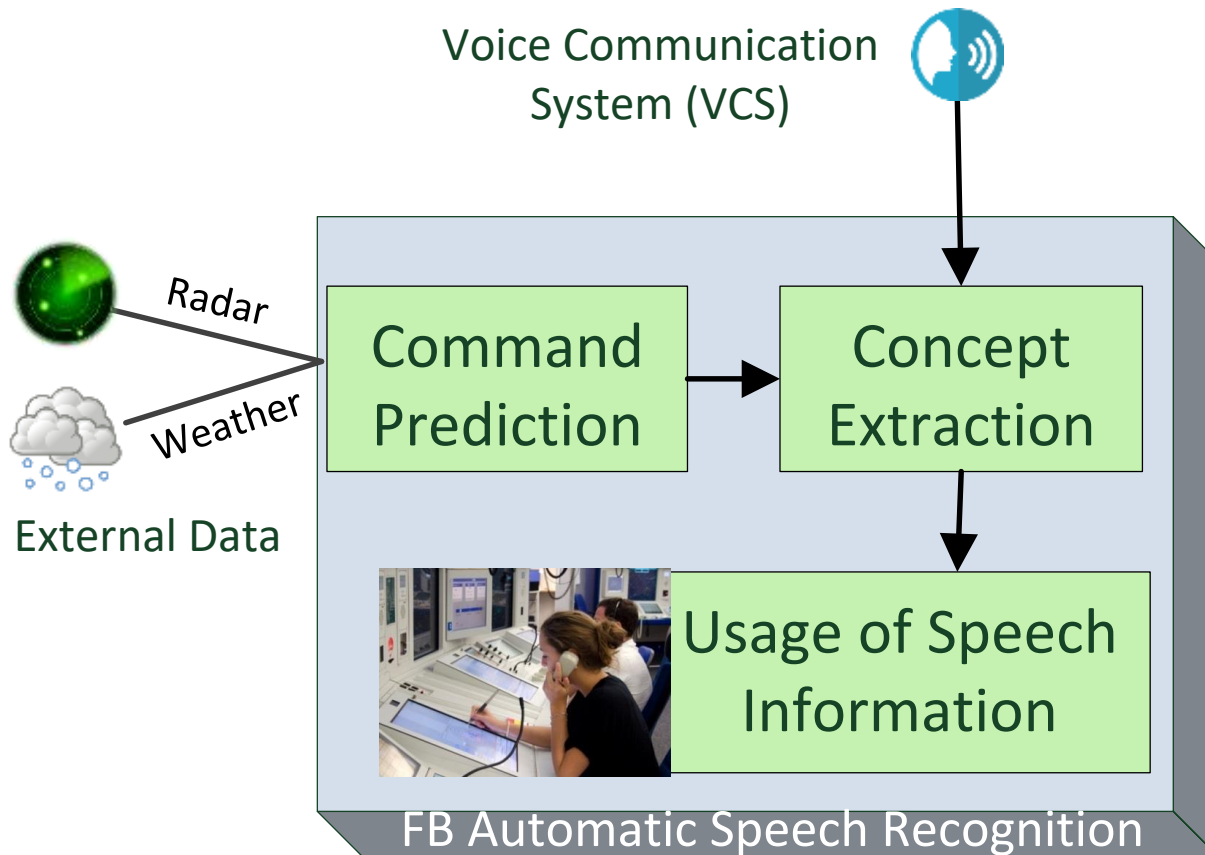
- E1: DFS: Callsign Highlighting + Consistency check with manual inputs
- E2: CRIDA / ENAIRE: Workload estimation
- E3: DLR / HC / ON: Command prediction for remote tower commands
- E4: Frequentis / DFS: Command Logging for Situation Awareness
- E5: DLR, Thales, Integra, Austro Control, ANS CR

From Paper to Electronic Flight Strips



- All flight information available in digital form in the system (and on the radar screen).
- Controllers have additional **workload**. Others have the benefits.

PJ.16-04: Common Achievements



- ASR as Functional Block (FB)
- 3 basic functions
- Ontology / Common Rules for Command Annotation

Contents

1. Automatic Speech Recognition
in SESAR context

2. Performed Exercise

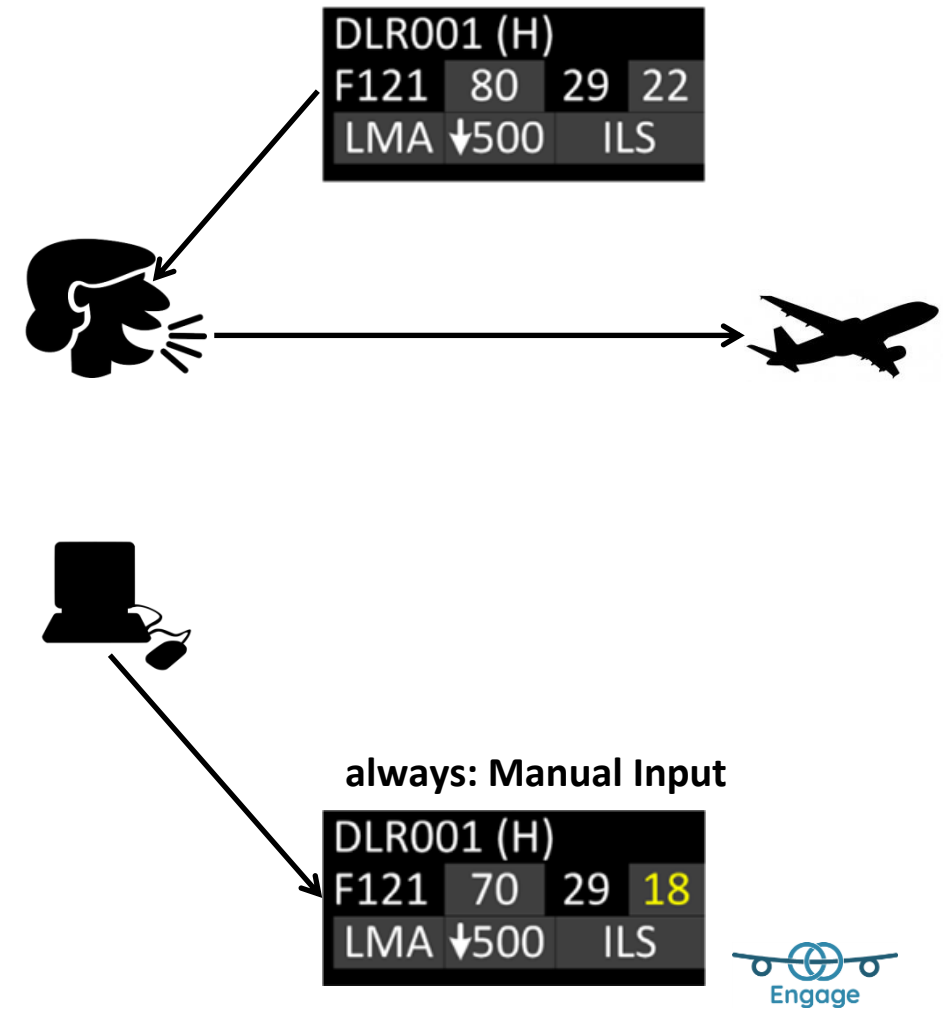
3. Exercise Results
 - for Prague Approach
 - for Vienna Approach
4. Conclusions



Baseline Scenario

Controller provides content of commands **twice**

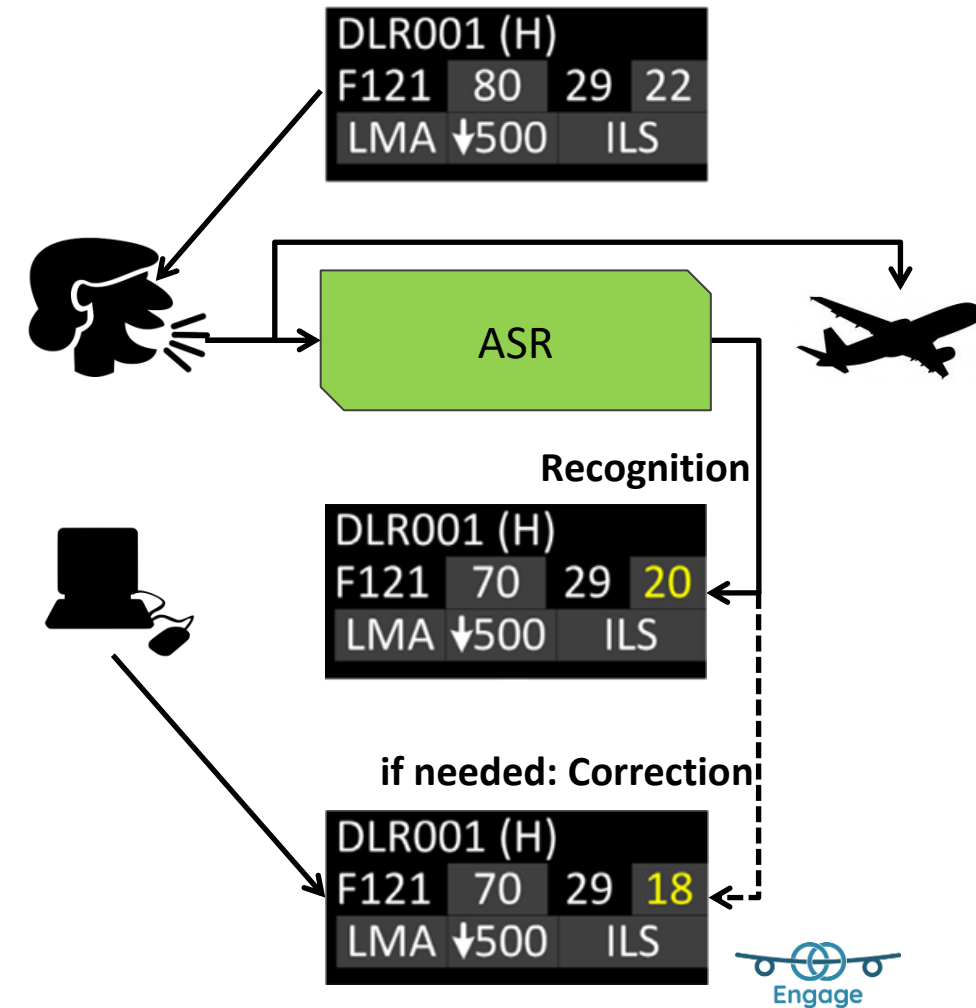
1. Verbal via voice to pilot
2. With mouse/keyboard to system



Solution Scenario

Automatic Speech Recognition (ASR)
instead of manual inputs

- Analyze controller-pilot communication
- Extract relevant commands
- Use extracted commands as label input
- Manual correction in case ASR fails



Performed Exercises and Setup

- February 2019 with 4 Prague Controllers (ANS-CR)
- May 2019 with 2 Vienna Controllers (ACG)
- Thales' Shape Platform
- Nuance 10.2.4 (Speech Server 6)
- Command Prediction from DLR
- Safety Assessment by Integra



*



**

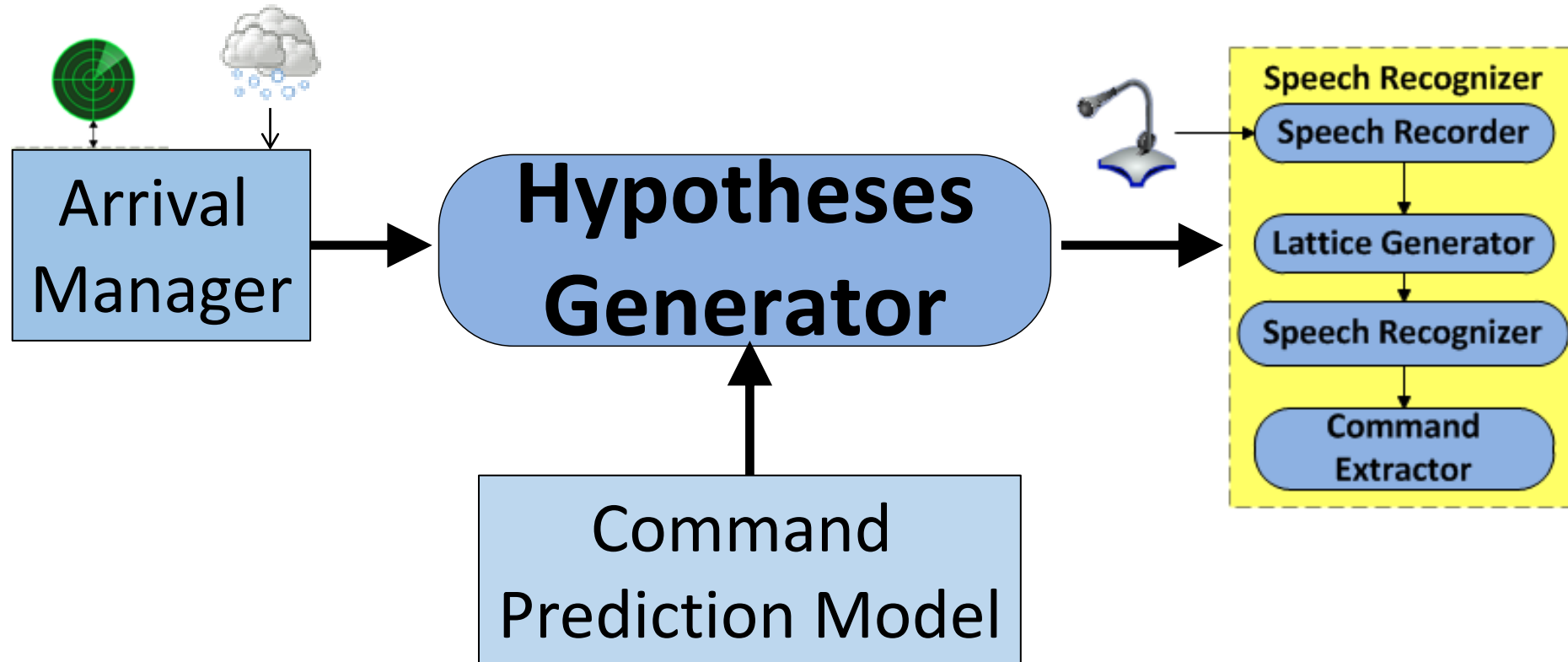
* <https://commons.wikimedia.org/w/index.php?curid=432452>
** <https://commons.wikimedia.org/w/index.php?curid=342954>



founding members



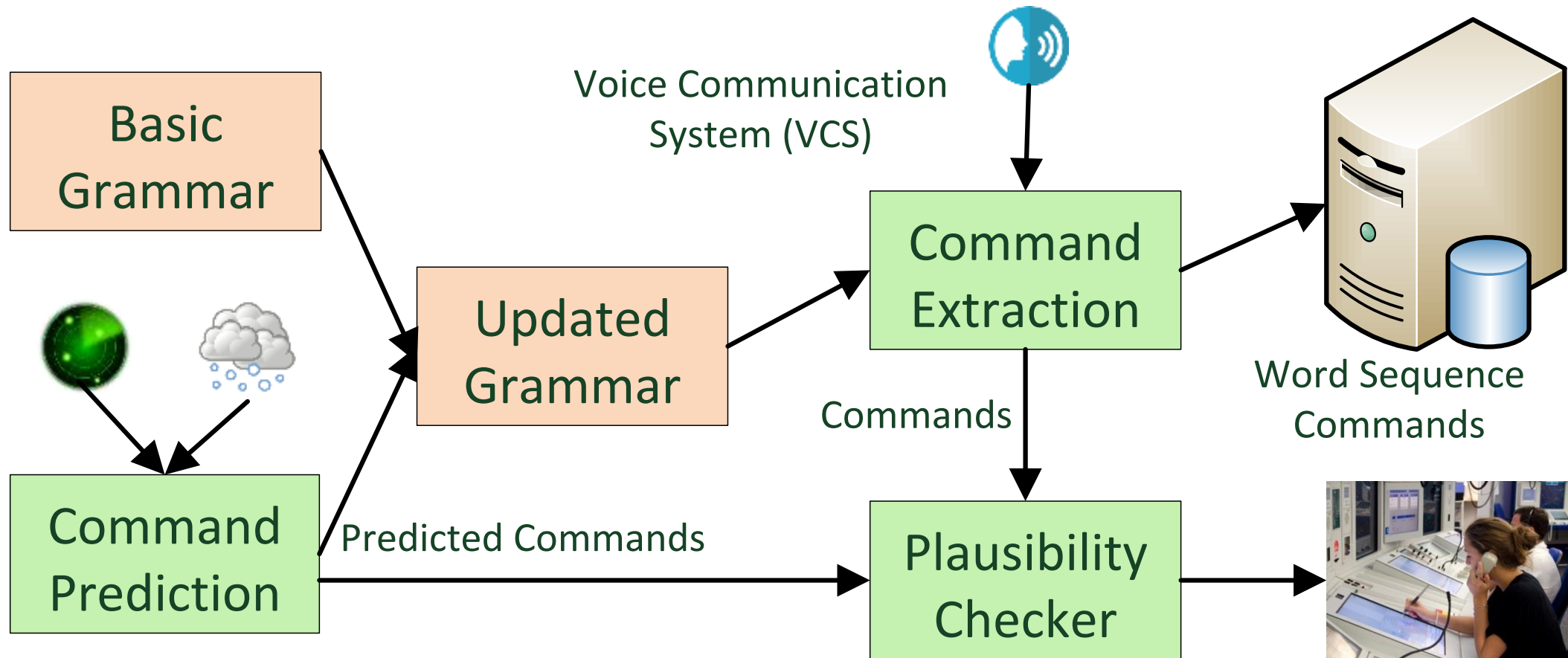
Assistant Based Speech Recognition (ABSR)



See paper:

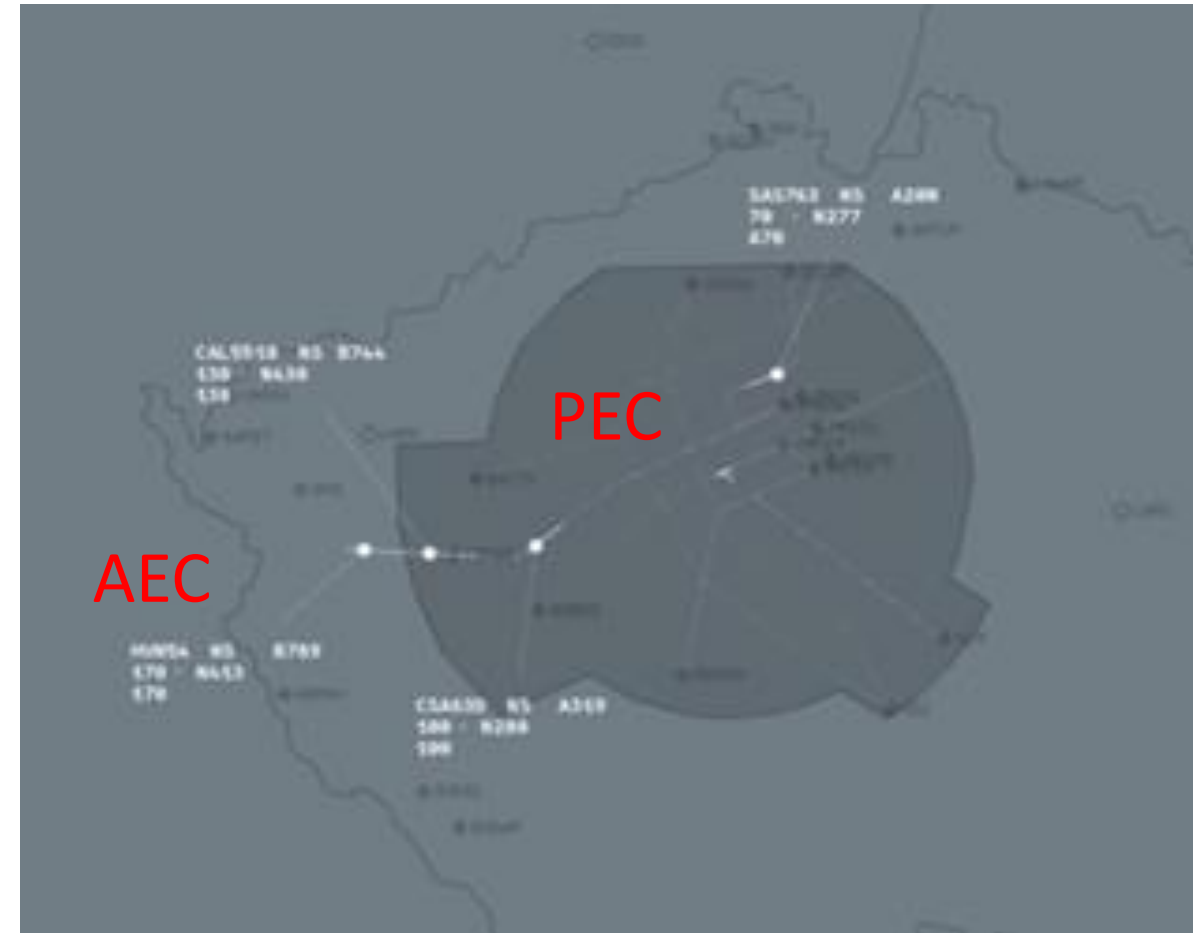
H. Helmke, J. Rataj, T. Mühlhausen, O. Ohneiser, H. Ehr, M. Kleinert, Y. Oualil, and M. Schulder, "Assistant-Based Speech Recognition for ATM Applications," in 11th USA/ Europe Air Traffic Management Research and Development Seminar (**ATM2015**), Lisbon, Portugal, 2015.

Implementation of ABSR in SESAR 2020 Exercise



Prague Exercise

- Inbounds plus outbounds
- No overflights
- Typical callsign mixes
- 24 (medium) resp. 30 (high traffic) movements
- AEC + PEC in one person
- AEC = Arrival Executive Controller
- PEC = Planning /Director Executive Controller
- Typical real-life scenario
→ reuse of MALORCA models



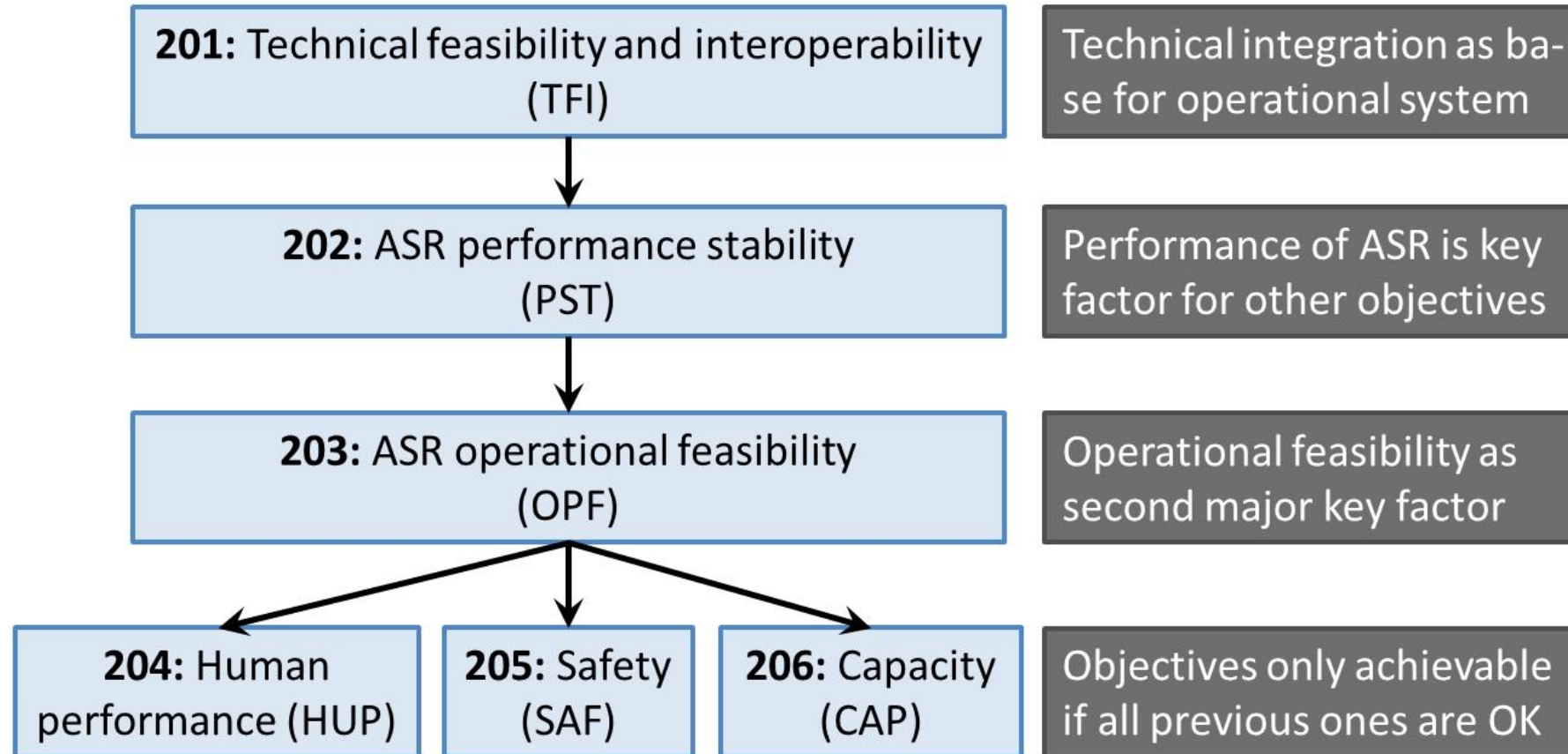
MALORCA = Machine Learning of Speech Recognition Models for Controller Assistance
See www.malorca-project.de
Exploratory Research Project of ER-2 call

Vienna Exercise

- Inbounds plus outbounds
- No overflights
- Typical callsign mixes
- 24 resp. 30 movements
- **3** Sectors plus Director for RWY34
- **No** typical real-life scenario
→ complete retraining of Command Prediction Model



Validation Objectives (of all PJ.16-04 Exercises)

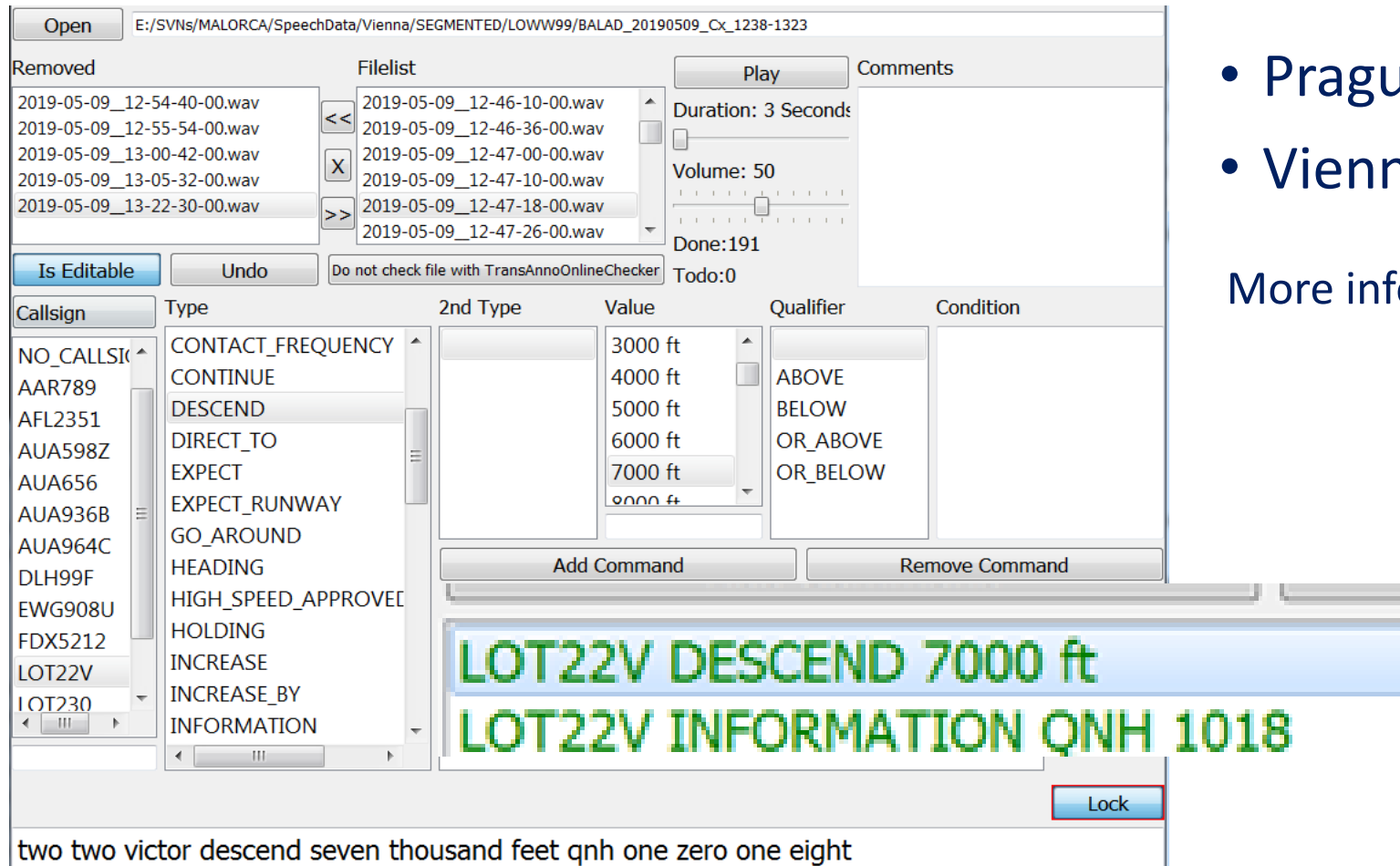


Contents

1. Automatic Speech Recognition in SESAR context
2. Performed Exercise
3. Exercise Results
 - **for Prague Approach**
 - for Vienna Approach
4. Conclusions



Tool Supported Command Annotation



The screenshot displays the SESAR tool interface for command annotation. The top section includes a file list and playback controls. The main area is a table for defining commands, with columns for Callsign, Type, 2nd Type, Value, Qualifier, and Condition. The bottom section shows the resulting command text and a lock button.

Callsign	Type	2nd Type	Value	Qualifier	Condition
NO_CALLSIGN	CONTACT_FREQUENCY		3000 ft		
AAR789	CONTINUE		4000 ft	ABOVE	
AFL2351	DESCEND		5000 ft	BELOW	
AUA598Z	DIRECT_TO		6000 ft	OR_ABOVE	
AUA656	EXPECT		7000 ft	OR_BELOW	
AUA936B	EXPECT_RUNWAY		8000 ft		
AUA964C	GO_AROUND				
DLH99F	HEADING				
EWG908U	HIGH_SPEED_APPROVED				
FDX5212	HOLDING				
LOT22V	INCREASE				
LOT230	INCREASE_BY				
	INFORMATION				

two two victor descend seven thousand feet qnh one zero one eight

- Prague: 16 runs (4036 commands)
- Vienna: 8 runs (2591 commands)

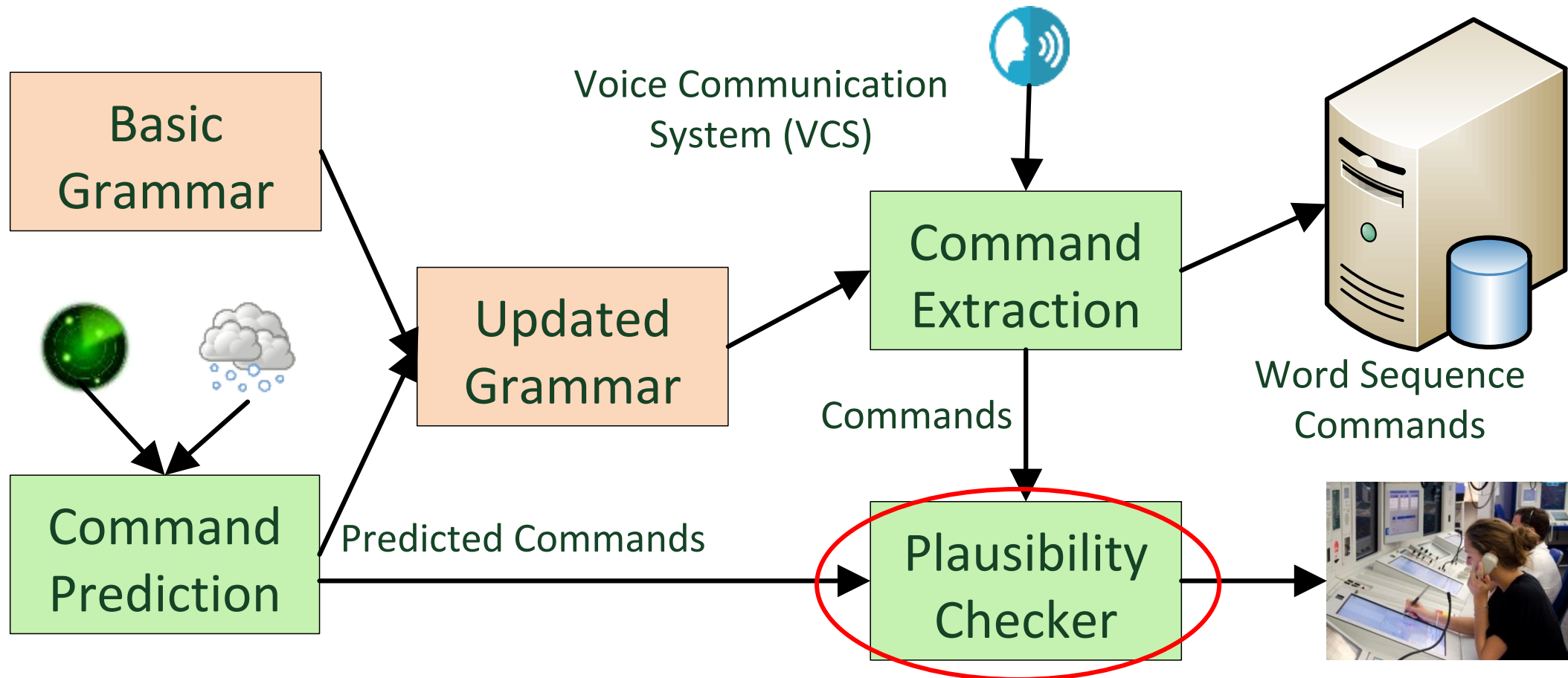
More information in next presentation

Prague ASR Performance without Plausibility Checker

	Total Given	RecognRate	ErrorRate	RejectRate
All commands	1994	70.41%	15.78%	14.75%
Must	1461	72.67%	17.23%	12.64%
Must + Should	1646	72.17%	15.72%	14.37%
> 10 commands	1989	70.69%	14.76%	15.34%

- Moderate recognition rates
- “Improvable” Error Rates

Exercise with Checker



Prague ABSR Performance with Checker

	Total Given	Without Checker		With Checker	
		Recogn Rate	Error Rate	Recogn Rate	Error Rate
All commands	1994	70.41%	15.78%	70.36%	5.36%
Must	1461	72.67%	17.23%	72.61%	5.32%

- No real decrease in recognition rate

Prague ABSR Performance with Checker

	Total Given	Without Checker		With Checker	
		Recogn Rate	Error Rate	Recogn Rate	Error Rate
All commands	1994	70.41%	15.78%	70.36%	5.36%
Must	1461	72.67%	17.23%	72.61%	5.32%

- No real decrease in recognition rate
- Significant Improvement of Error Rate

Prague Command Prediction Performance

	Total Given	Recogn Rate	Error Rate	# CmdE	CmdPER
All commands	1994	70.36%	5.36%	6	0.31%
Must	1461	72.61%	5.32%	2	0.15%
Must + Should	1646	72.11%	4.84%	2	0.13%
> 10 commands	1989	70.65%	5.32%	2	0.11%

- Very good command prediction accuracy
- Models trained on ops room data are reusable in lab environment

CmdE = Absolute Number of Command Prediction Errors
CmdPER = Command Prediction Error Rate in Percent

Contents

1. Automatic Speech Recognition in SESAR context
2. Performed Exercise
3. Exercise Results
 - for Prague Approach
 - **for Vienna Approach**
4. Conclusions



Vienna A^BSR Performance

	Total Given	Without Checker		With Checker	
		Recogn Rate	Error Rate	Recogn Rate	Error Rate
All commands	1259	54.77%	20.39%	52.28%	6.14%
Must	1094	54.42%	22.07%	53.29%	6.56%
Must + Should	1094	54.34%	22.16%	53.22%	6.65%
> 10 commands	1251	55.45%	16.91%	52.93%	5.84%

- Slight decrease in recognition rate
 - Again **Significant Improvement of Error Rate**

Vienna Command Prediction Performance

	Total Given	Recogn Rate	Error Rate	# CmdE	CmdPER
All commands	1259	52.28%	6.14%	58	4.77%
Must	1094	53.29%	6.56%	27	2.56%
Must + Should	1094	53.22%	6.65%	27	2.56%
> 10 commands	1251	52.93%	5.84%	53	4.33%

- Acceptable command prediction accuracy
- Model Retraining might have an effect on performance

CmdE = Absolute Number of Command Prediction Errors
CmdPER = Command Prediction Error Rate in Percent

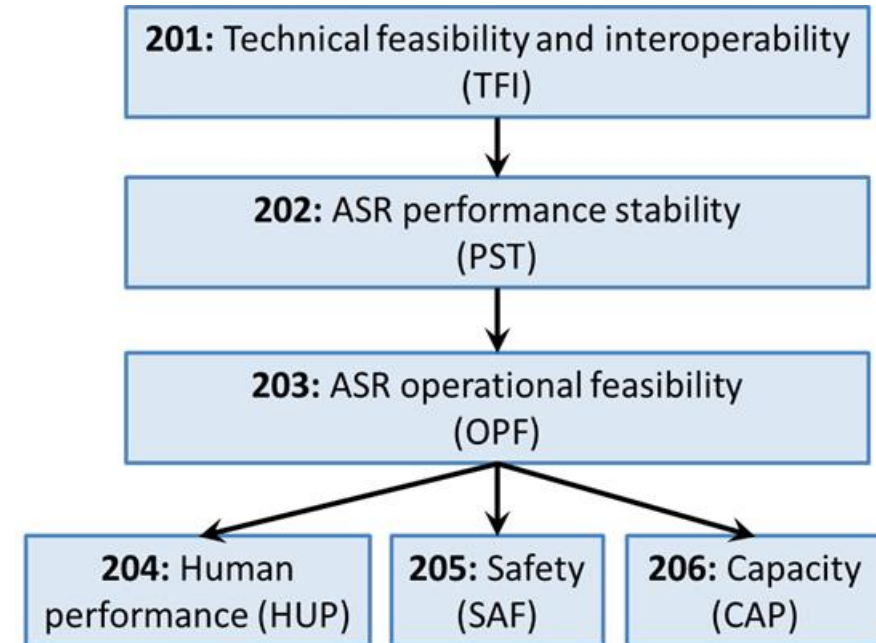
Command Recognition Rates per Controller

Prague	Both Scenarios	Vienna	Both Scenarios
ATCO 1	73.7%		
ATCO 2	81.7%	ATCO 11	78.2%
ATCO 3	78.3%	ATCO 12	52.5%
ATCO 4	79.0%		

- Problems with one Vienna controller
(Language Model versus phraseology)

Results for Validation Objectives

- Technical feasibility ✓
- ASR performance stability ✓
- ASR operational feasibility 😞
- Human performance, safety, capacity? ⚠️



Challenges of Usage of COTS ASR Engines

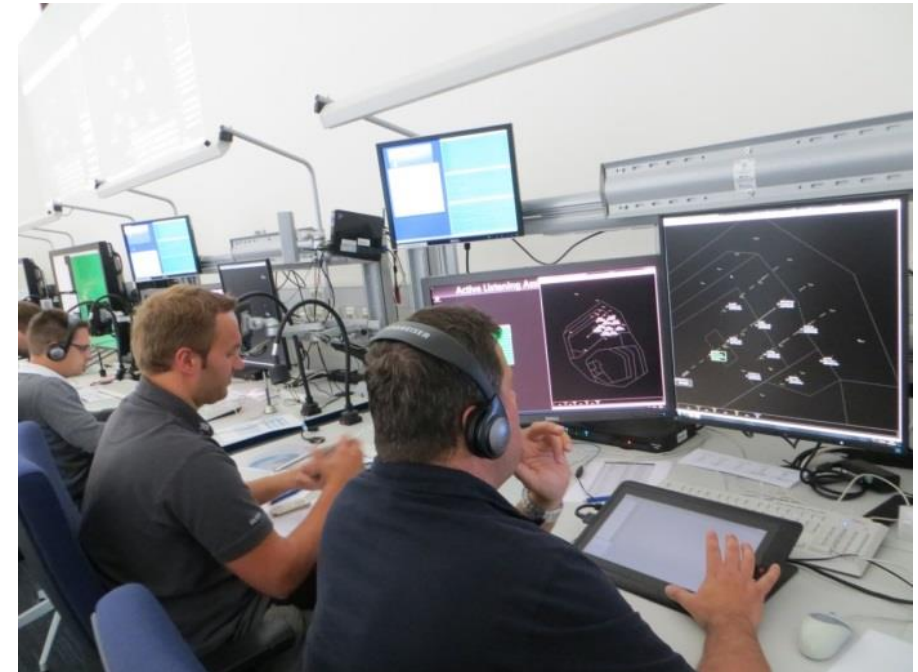
- Cloud is no option
 - Data privacy requirements
 - Cyber security challenges
 - Safety challenges
 - Real-time aspects are important
- ATC has a special phraseology
- Context information is available
 - ATC Language model / ATC grammar
 - Special acoustic models for accents

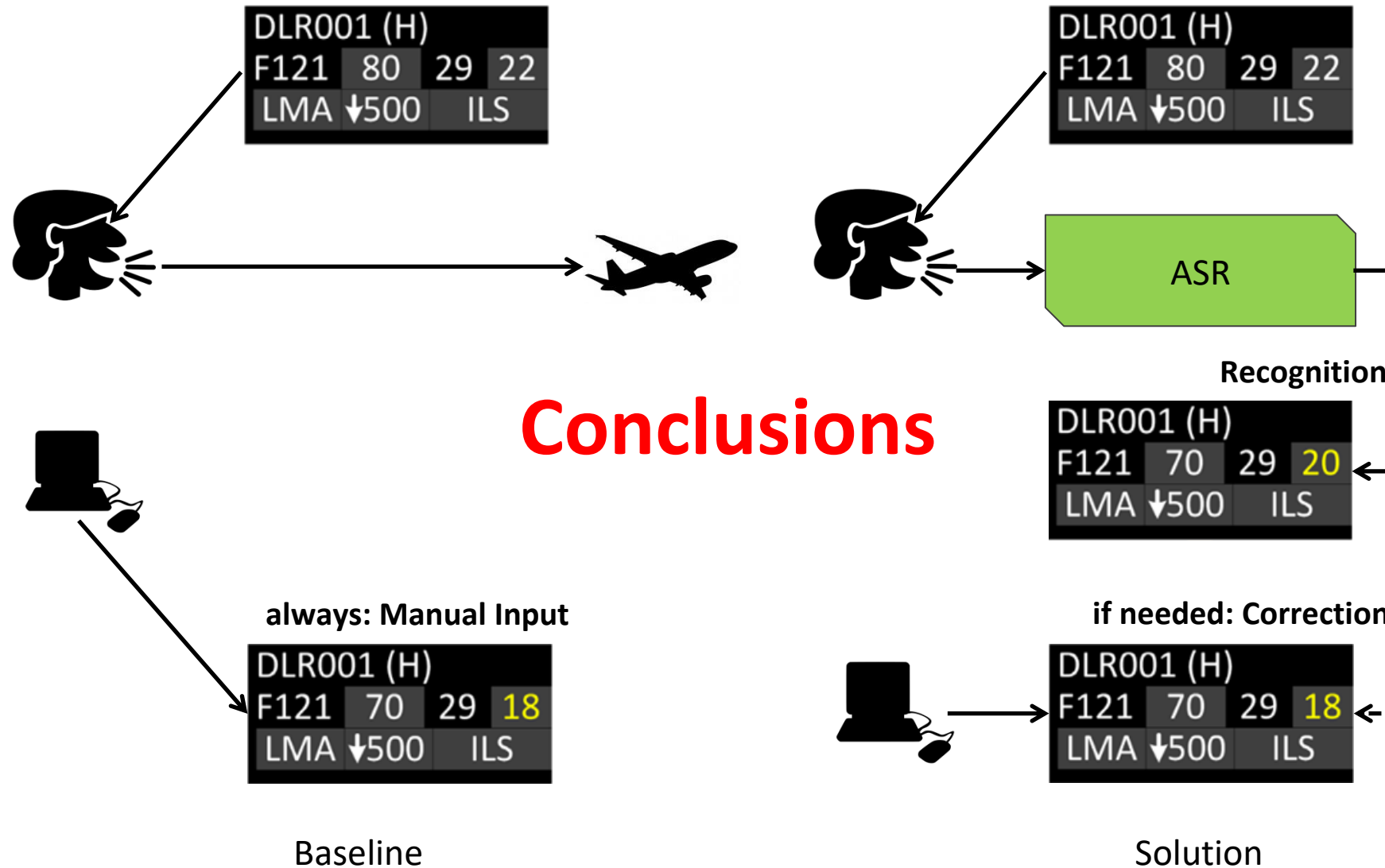


Contents

1. Automatic Speech Recognition in SESAR context
2. Performed Exercise
3. Exercise Results
 - for Prague Approach
 - for Vienna Approach

4. Conclusions







Conclusions of Performed Exercise

- Used COTS approach results in recognition rates of [31% .. 82%]
- A^BSR with checker based on **command prediction** significantly reduces command error rates:
from [14.7% .. 22.6%] to [4.8% .. 6.6]
- Most of the false recognitions are not shown to the ATCo
- HMI with A^BSR assisted radar label maintenance re-validated
- Still, no idea of needed ASR performance

Recommendations for SESAR Wave #2

- Faster recognition times
- 1:1 usage of COTS not “recommended”
 - Special acoustic model trained for final end users
 - Extend grammar to support more phraseology deviations
 - Train on real life data
- Use automatic transcriptions for model training
- More validation runs for statistical significance needed
- Use of common ontology for data exchange



THALES



Air Navigation Services
of the Czech Republic

Covering the sky...



Thank you very much for your attention!



This project has received funding from the SESAR Joint Undertaking under the European Union's Horizon 2020 research and innovation programme under grant agreement No 734141.



CWP HMI



Founding Members



EUROPEAN UNION



EUROCONTROL