MALORCA
Machine Learning of Speech Recognition Models for Controller Assistance

DLR | UNIVERSITÄT DES SAARLANDES | idiap RESEARCH INSTITUTE | austro CONTROL | Air Navigation Services of the Czech Republic

Covering the sky...

Building Blocks of Assistant Based Speech Recognition for Air Traffic Management Applications

Matthias Kleinert, Hartmut Helmke, Heiko Ehr, Gerald Siol (DLR)
Christian Kern (Austro Control)
Dietrich Klakow, Mittul Singh (UdS)
Petr Motlicek (Idiap)

German Aerospace Center (DLR), Braunschweig
Saarland University (UdS), Saarbrücken
Idiap Research Institute, Martigny
Austro Control, Vienna

Founding Members
EUROPEAN UNION EUROCONTROL

Content

- **Motivation for Speech Recognition in Air Traffic Control**
- Building Blocks of Assistant Based Speech Recognition (ABSR)
- How to adapt ABSR to new environments
- Influence of different components of ABSR on the performance
- Conclusions

MALORCA

From Paper to Electronic Flight Strips

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

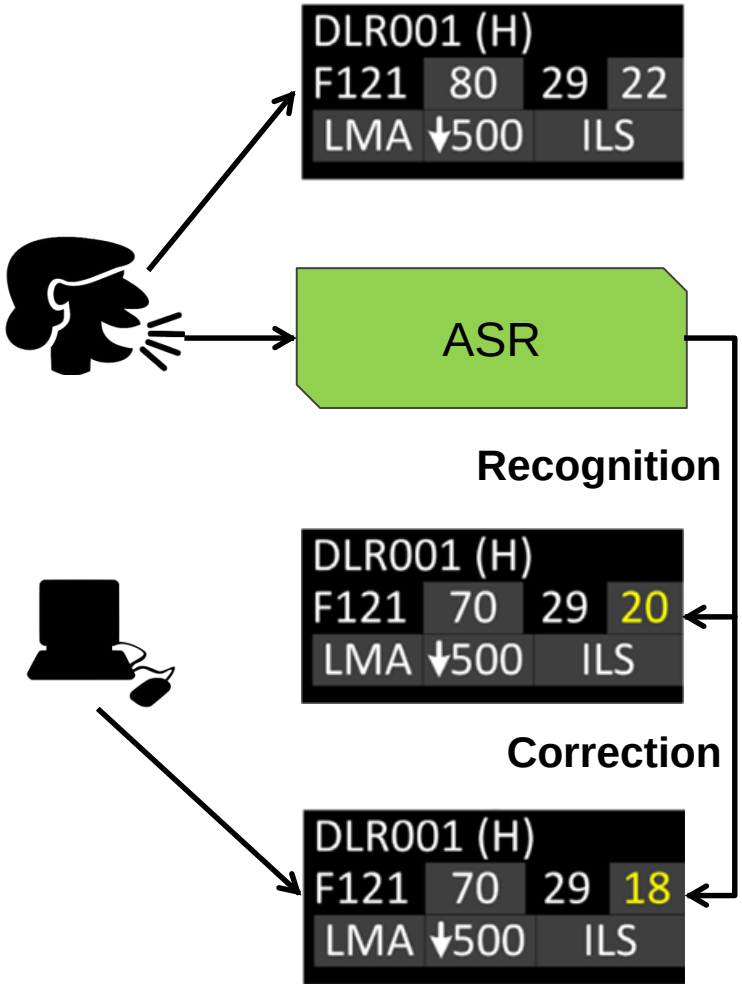
DLR

MALORCA

How can Speech Recognition help?

- Automatic Speech Recognition (ASR) can be the solution
  - Analyze controller-pilot communication
  - Extract relevant commands
  - Use commands as label input
  - Manual correction in case ASR fails

- For this DLR and Saarland University developed Assistant Based Speech Recognition (ABSR) to achieve:
  - →High recognition and low error rates on command level
  - →Reduction of controller workload by automating flight strip management



DLR

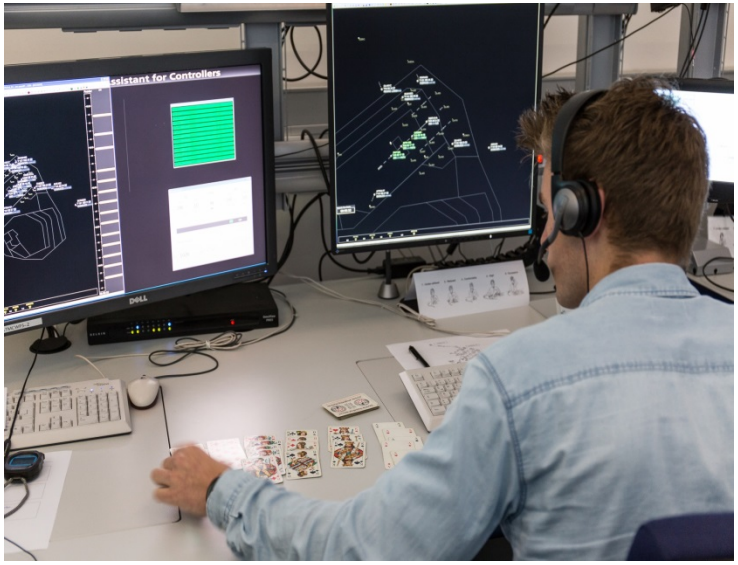
ABSR-Validation at DLR in 2015

- In 2014/2015 > 20 controllers from DFS, Austro Control and ANS CR validated the ABSR system developed by UdS and DLR for Dusseldorf Approach Area.

- **Goals** → Quantify the benefits for ABSR with respect to
  - controllers' workload and
  - ATM efficiency.

- **Baseline:**
  - Commands entered by mouse into radar labels

- **Improved Mode:**
  - Commands entered by ABSR correction if necessary by mouse



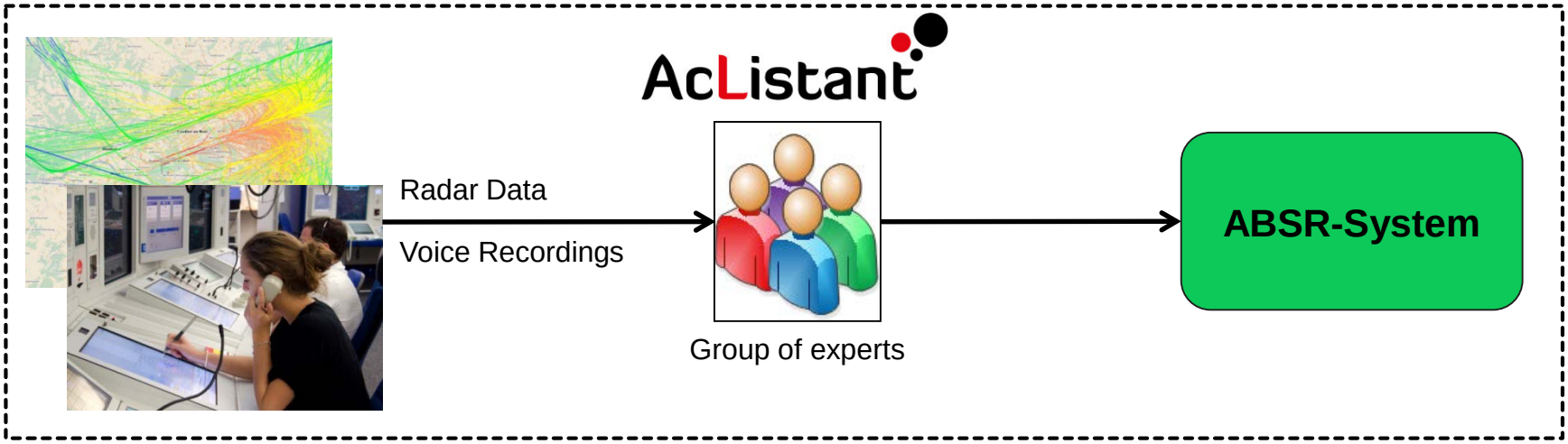
ABSR-Results of 2015

- **Performance**
  - → Recognition rate: 95%
  - → Error rate: 2%
- **ANSPs and controllers**
  - → Radar Label maintenance time reduced by a factor of 3
- **Airports**
  - → benefit from increase flow of 1 to 2 landings per hour **Performance**
- **Airlines**
  - → save 50 to 65 liters of kerosene per flight
- **Society**
  - → saves approx. 130 kg of CO2 per flight*

* A320, 0.8 kg / l, 1 kg kerosene results in 3.15 kg C02; 35 landings per hour extrapolation of results of 60 minutes scenarios for 23R, 8 controllers…, see papers at DASC 2016 and FAA/Eurocontrol ATM Seminar 2017

MALORCA

What is the Problem?

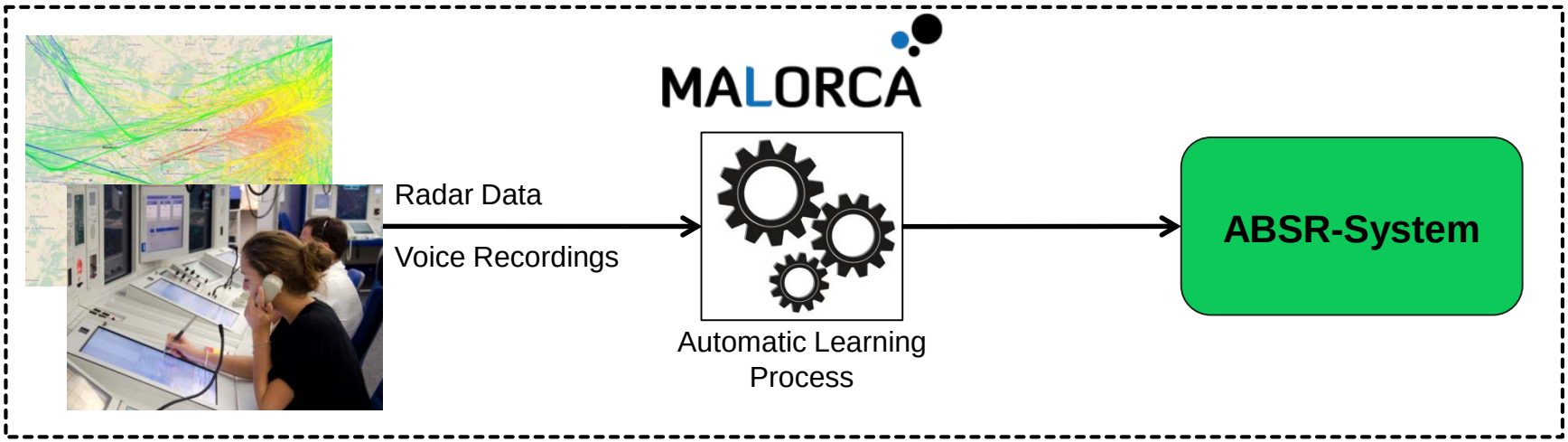


- In AcListant® 1.3M Euro were spent for ABSR development and validation just for Düsseldorf Approach
- Adaptation is necessary for other airports and working positions because of:
  - Different accents, working procedures, airspace layouts etc.
- Requires lots of time and expert knowledge
- Adaptation will become necessary again even in an already deployed system due to changes in working procedures, airspace layouts etc. (maintenance)

DLR

MALORCA

How to solve it?



- Instead of (highly skilled and paid) experts, machine learning is used.
- Horizon 2020 funded project MALORCA (Machine Learning of Speech Recognition Models for Controller Assistance)

DLR idiap RESEARCH INSTITUTE austro CONTROL Air Navigation Services of the Czech Republic Covering the sky... UNIVERSITÄT DES SAARLANDES

DLR

Content

- Motivation for Speech Recognition in Air Traffic Control
- **Building Blocks of Assistant Based Speech Recognition (ABSR)**
- How to adapt ABSR to new environments
- Influence of different components of ABSR on the performance
- Conclusions

MALORCA

From Speech Signal to HMI

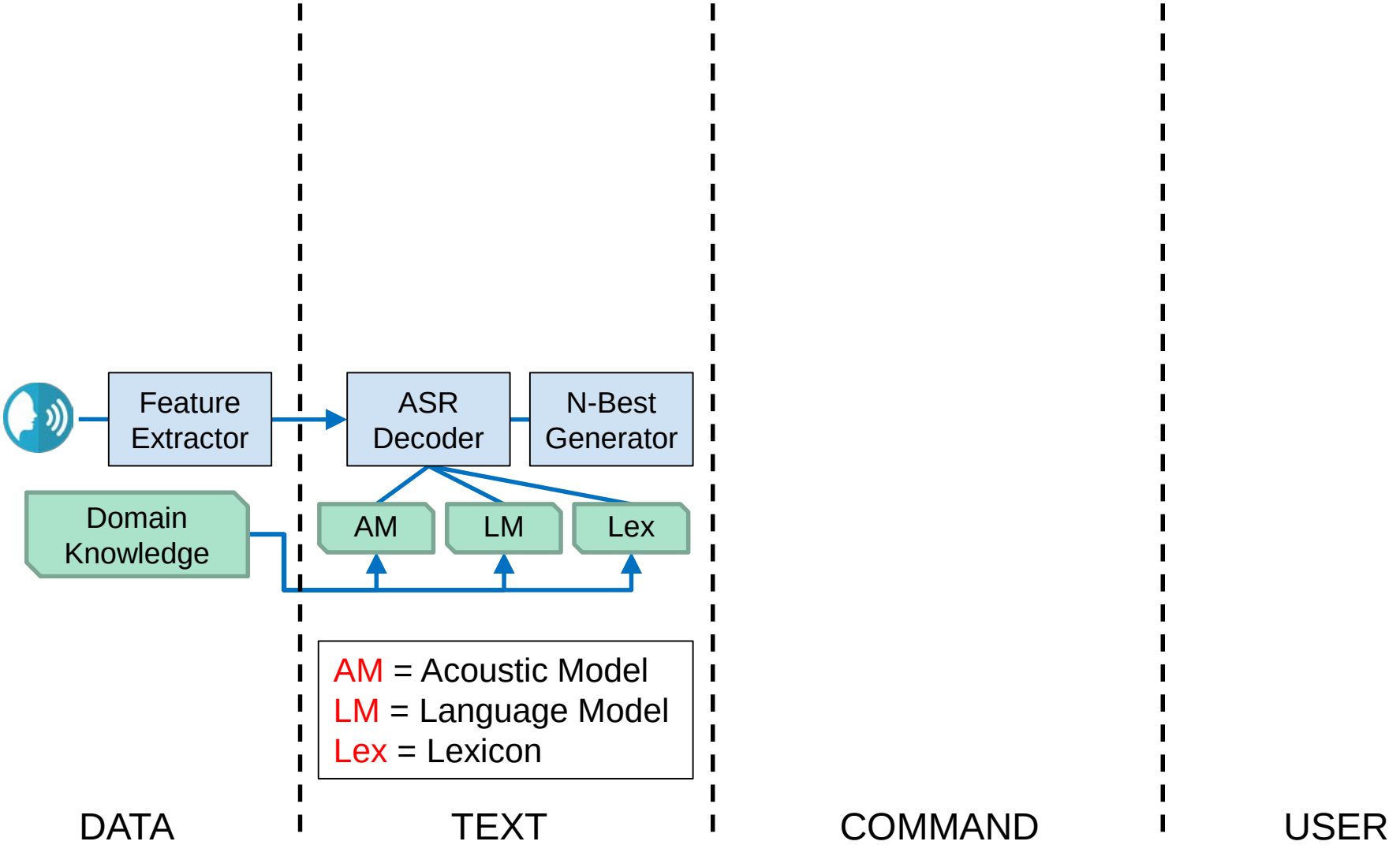
Feature Extractor

DATA | TEXT | COMMAND | USER

DLR

MALORCA

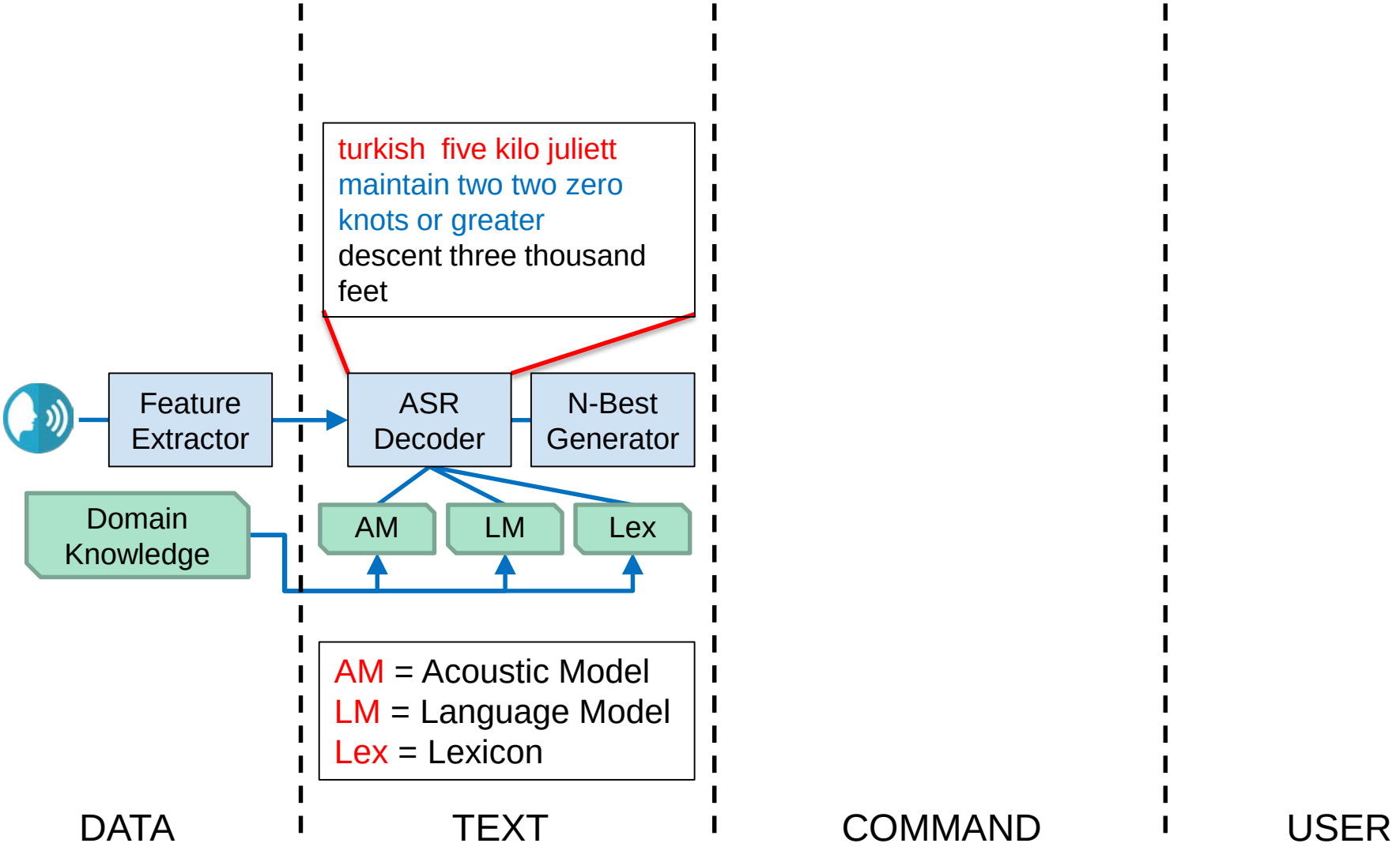
From Speech Signal to HMI



DLR

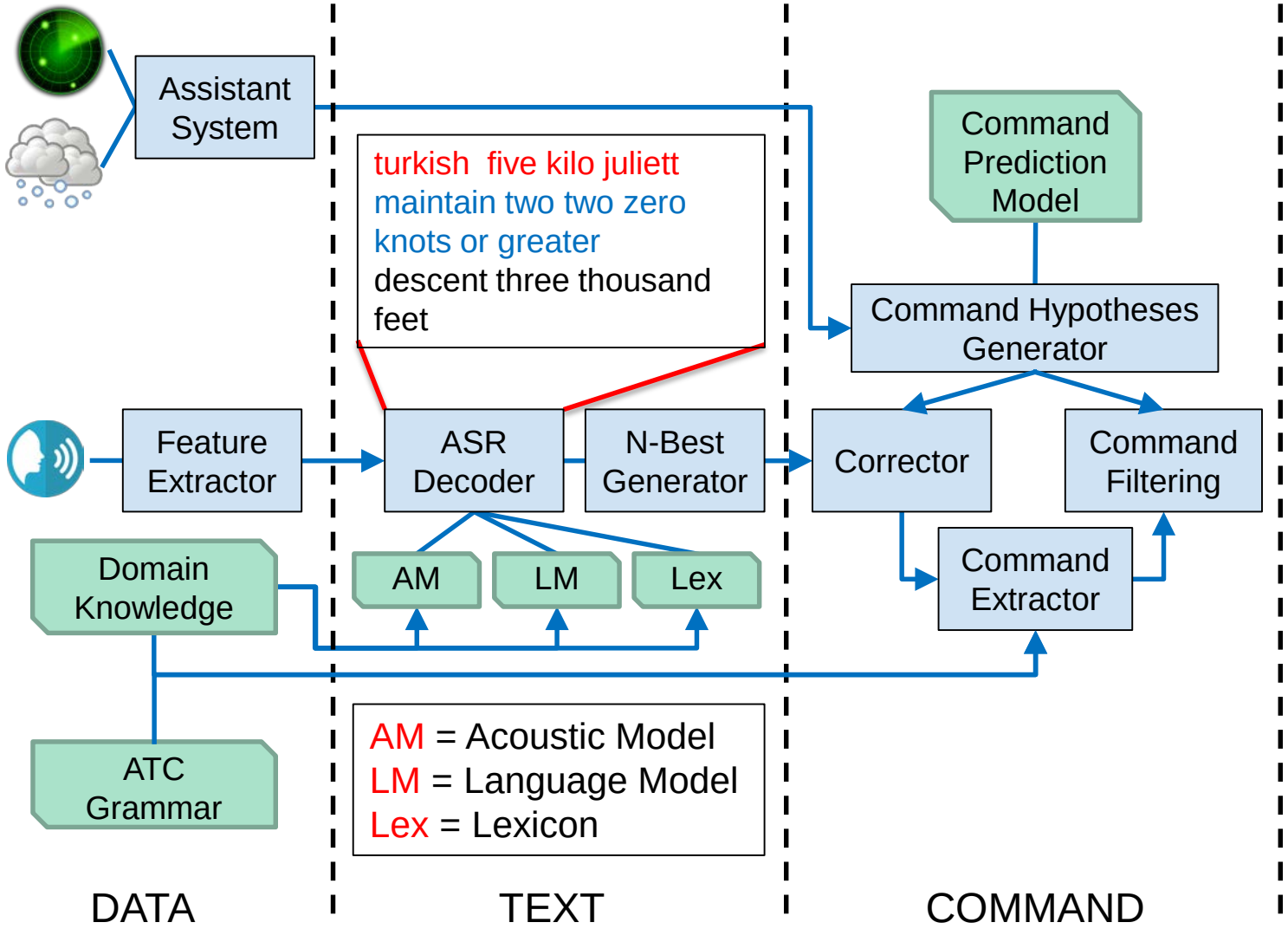
MALORCA

From Speech Signal to HMI



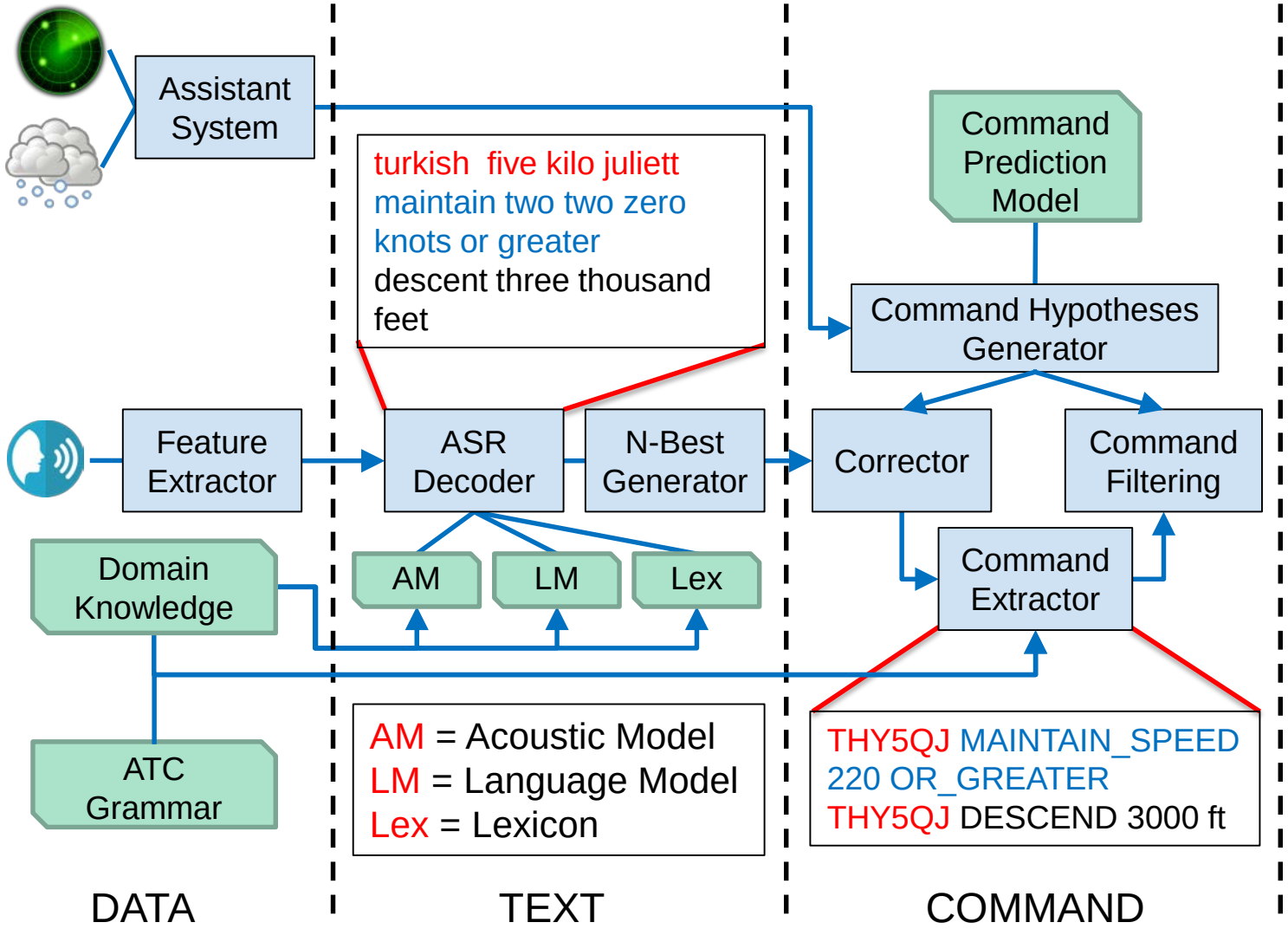
MALORCA

From Speech Signal to HMI



MALORCA

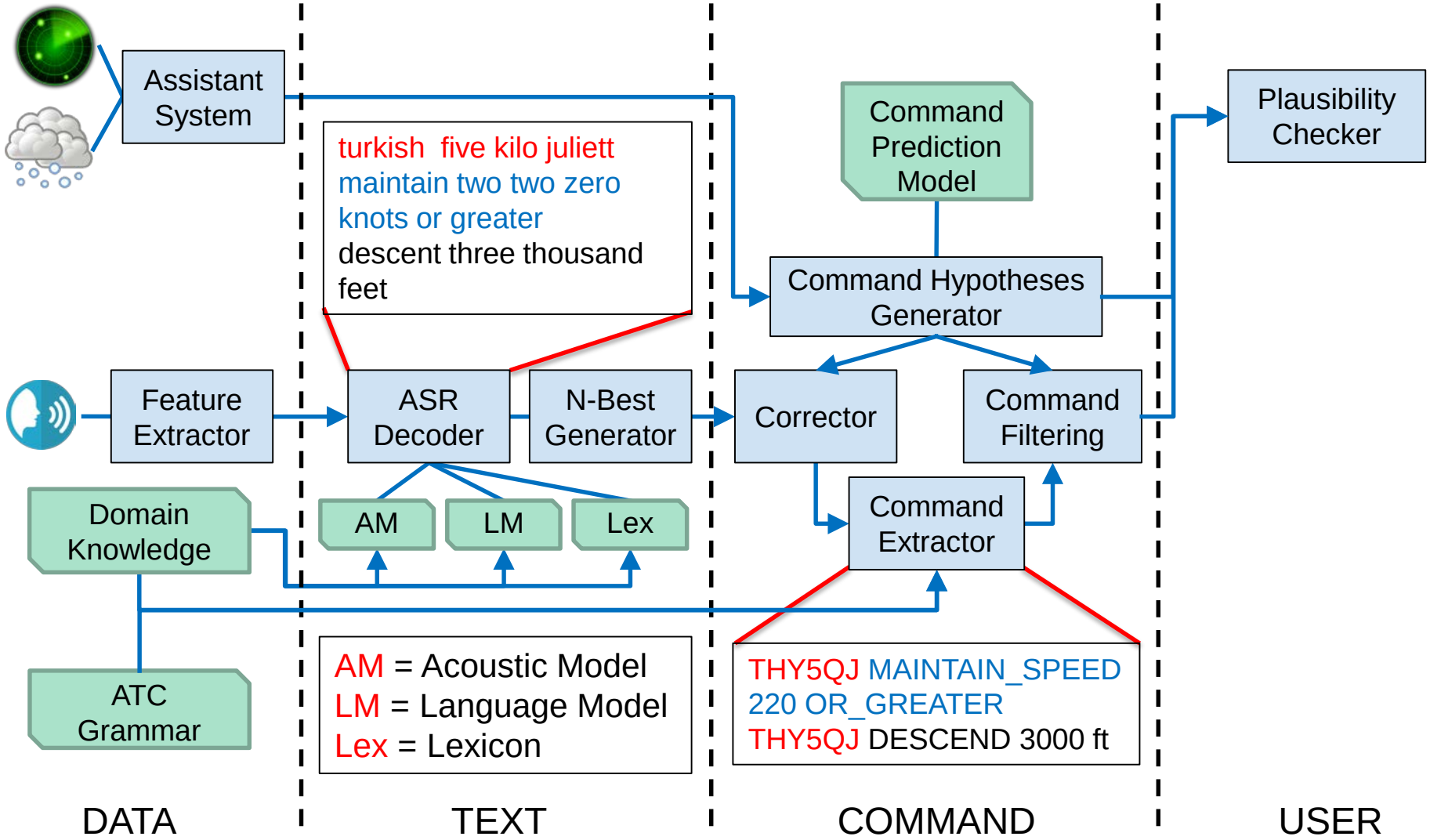
From Speech Signal to HMI



DLR

MALORCA

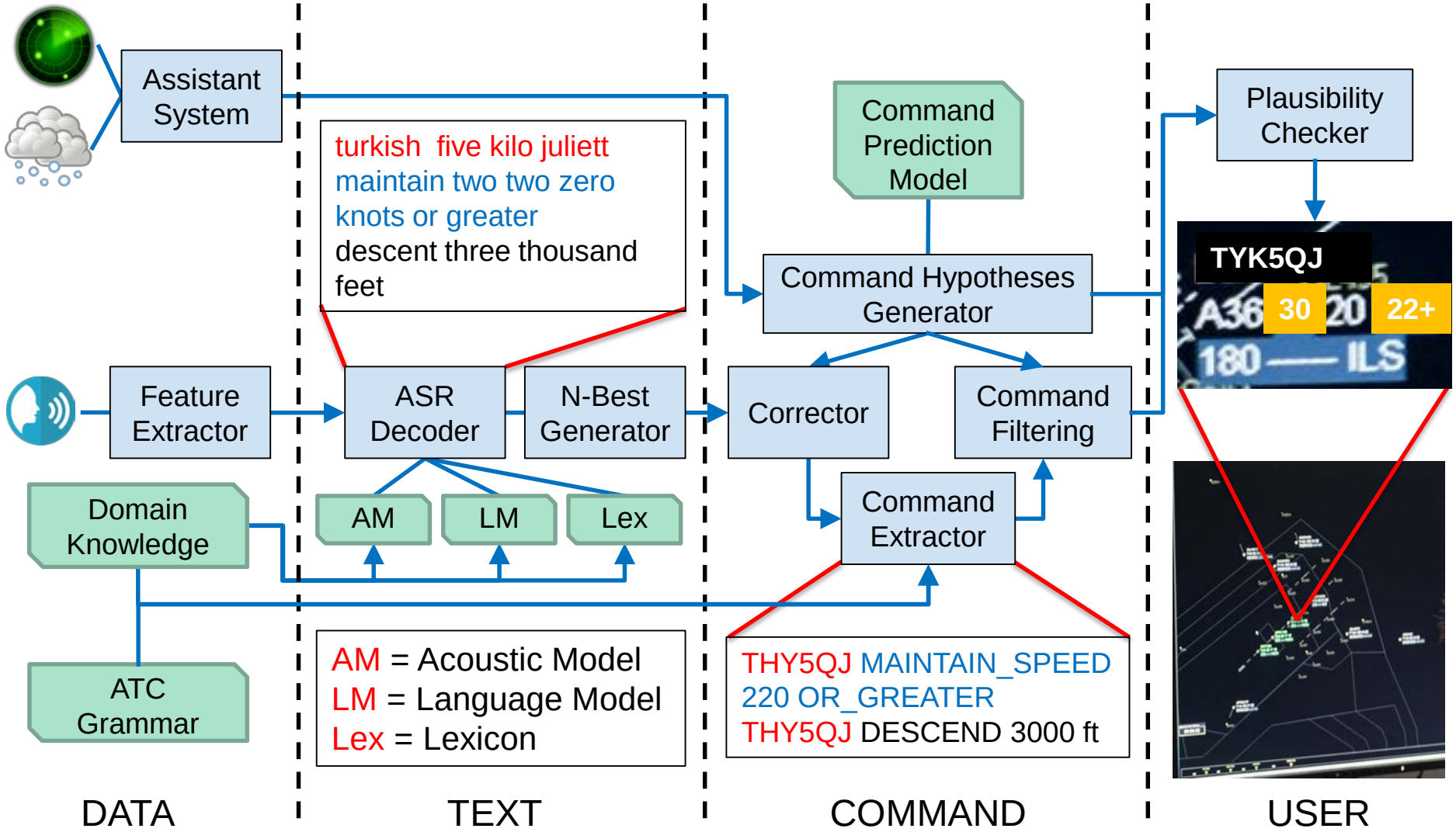
From Speech Signal to HMI



DLR

MALORCA

From Speech Signal to HMI

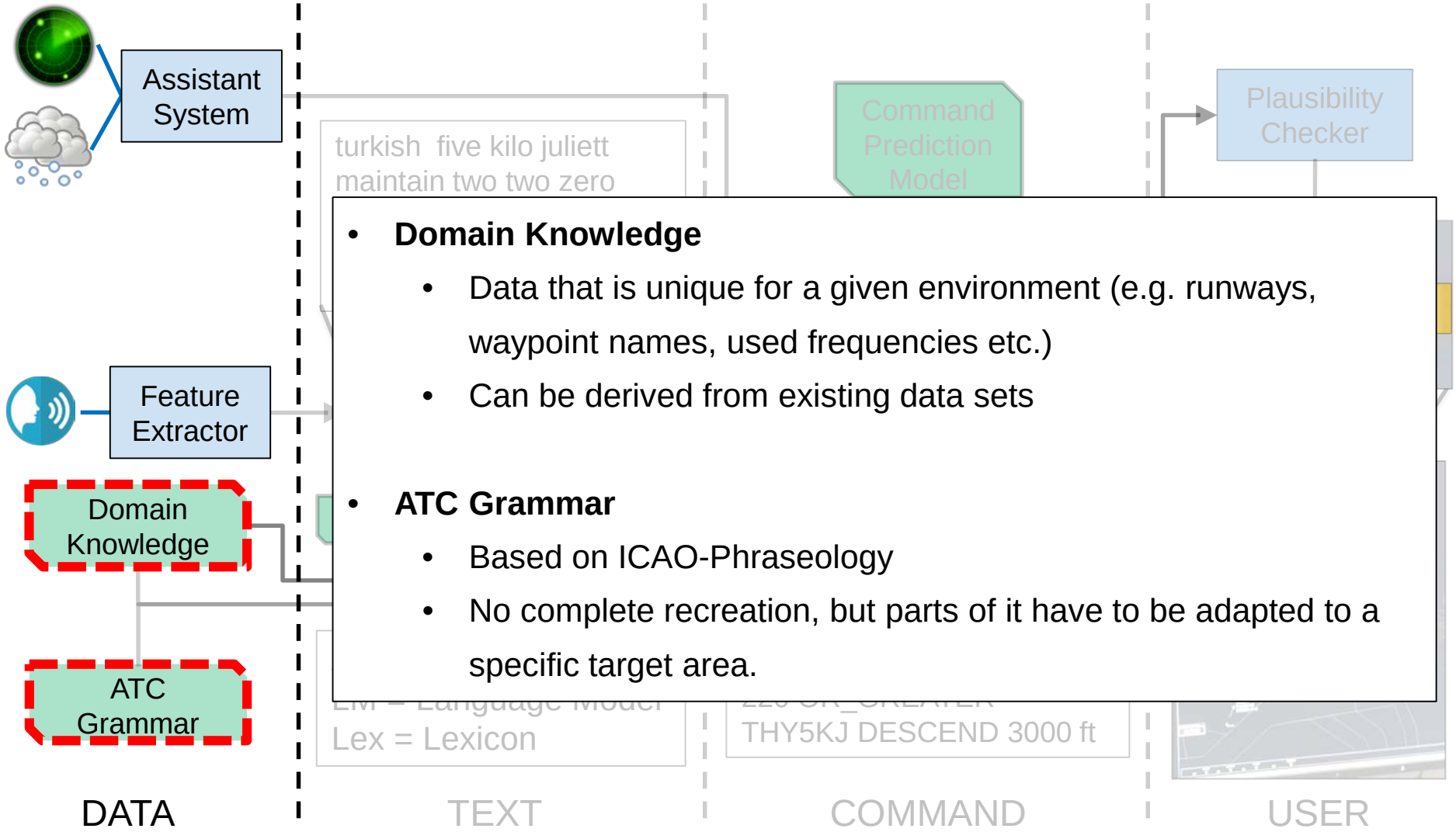


Content

- Motivation for Speech Recognition in Air Traffic Control
- Building Blocks of Assistant Based Speech Recognition (ABSR)
- **How to adapt ABSR to new environments**
- Influence of different components of ABSR on the performance
- Conclusions

MALORCA

Adaptations to DATA-Module

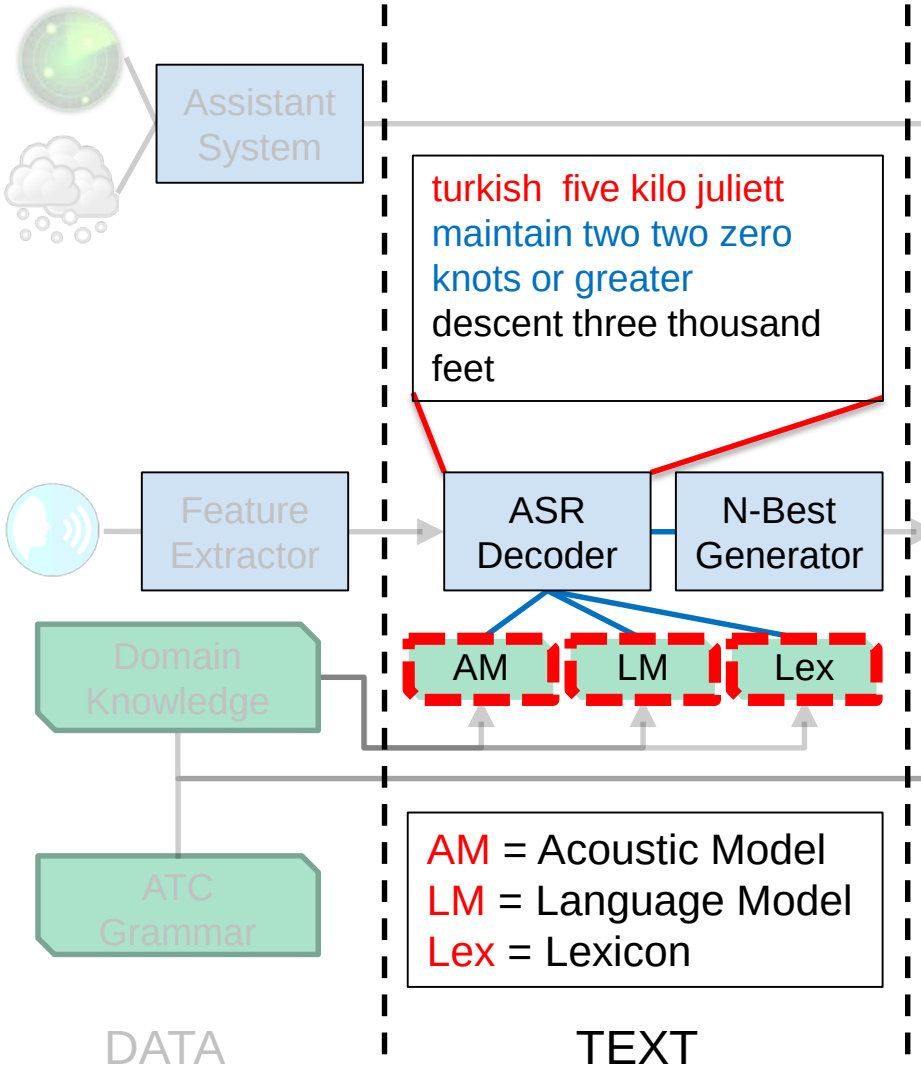


- **Domain Knowledge**
  - Data that is unique for a given environment (e.g. runways, waypoint names, used frequencies etc.)
  - Can be derived from existing data sets
- **ATC Grammar**
  - Based on ICAO-Phraseology
  - No complete recreation, but parts of it have to be adapted to a specific target area.

DLR

MALORCA

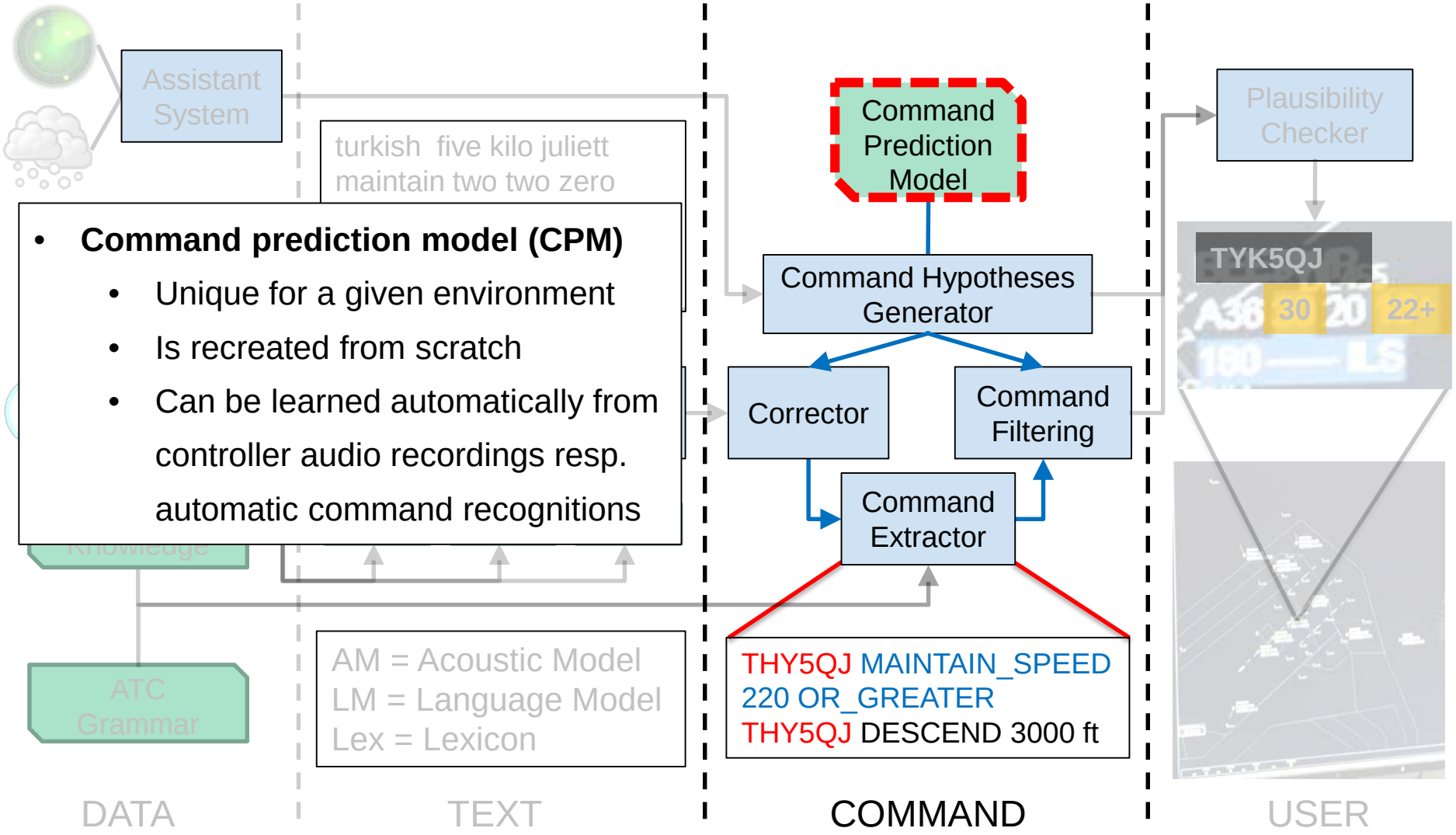
Adaptations to TEXT-Module



- **Acoustic model**
  - Trained from transcribed and/or untranscribed data
  - Speaker-dependent models possible
- **Language model**
  - Learned from labeled ATC text transcriptions
  - Grammar can be used for automatic labeling of training data
  - Data from existing language models can be reused
- **Lexicon**
  - Local Words + pronunciations for a specific environment (e.g. Greetings)

MALORCA

Adaptations to COMMAND-Module

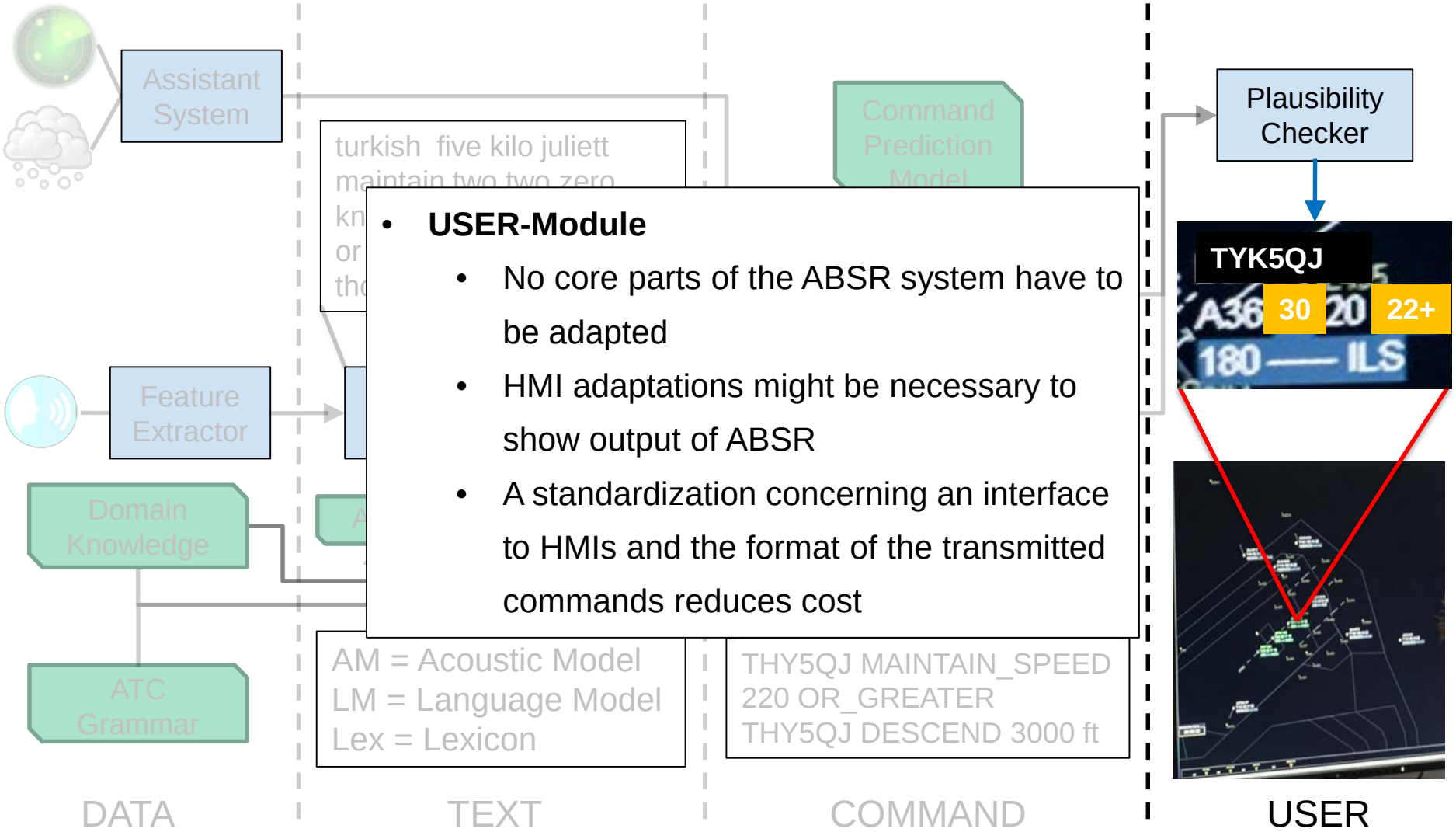


- **Command prediction model (CPM)**
  - Unique for a given environment
  - Is recreated from scratch
  - Can be learned automatically from controller audio recordings resp. automatic command recognitions

DLR

MALORCA

Adaptations to USER-Module



- **USER-Module**
  - No core parts of the ABSR system have to be adapted
  - HMI adaptations might be necessary to show output of ABSR
  - A standardization concerning an interface to HMIs and the format of the transmitted commands reduces cost

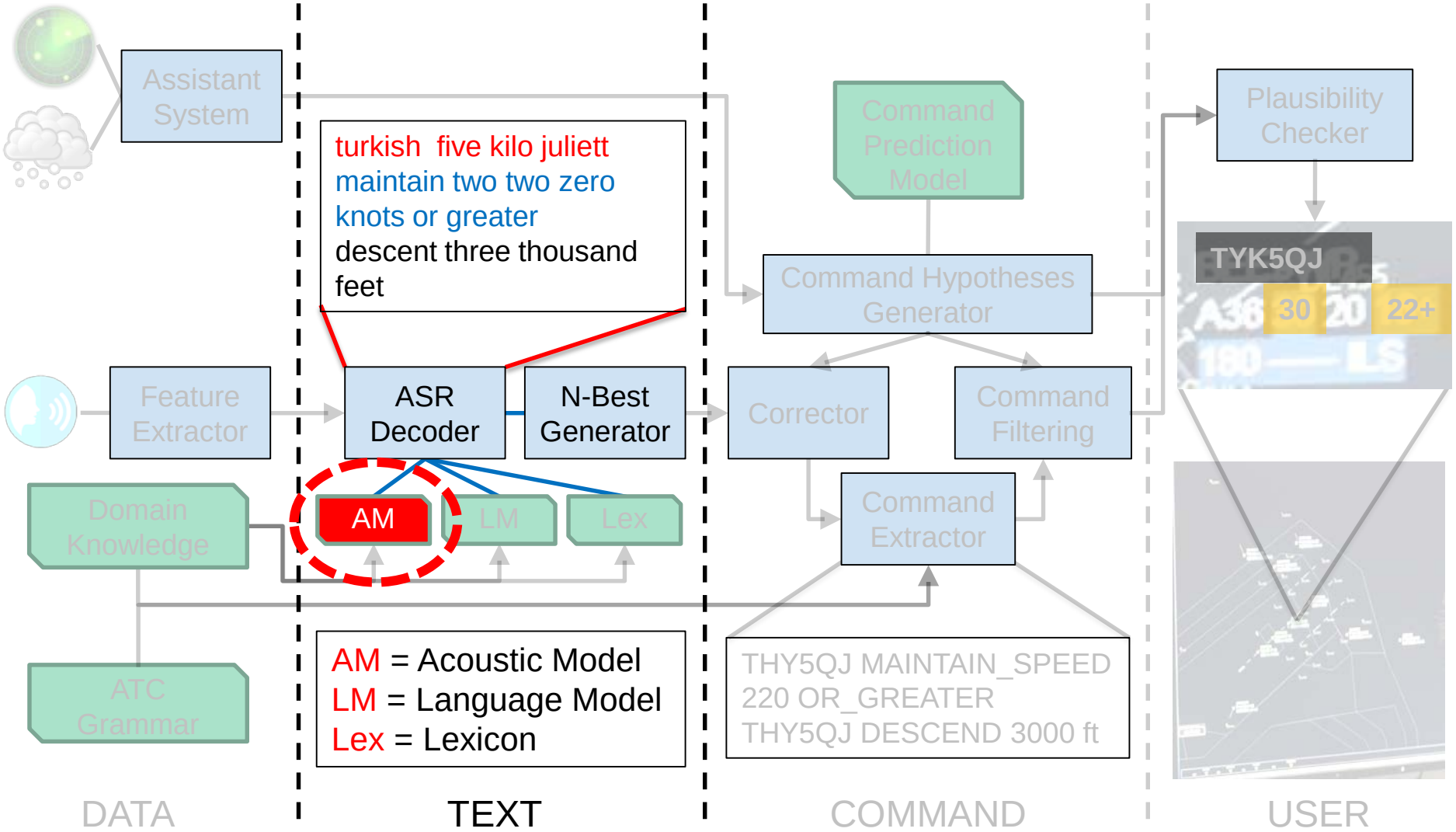
DLR

Content

- Motivation for Speech Recognition in Air Traffic Control
- Building Blocks of Assistant Based Speech Recognition (ABSR)
- How to adapt ABSR to new environments
- **Influence of different components of ABSR on the performance**
- Conclusions

MALORCA

Speaker-Dependent vs Speaker-Independent Acoustic Model



DLR

MALORCA

Speaker-Dependent vs Speaker-Independent Acoustic Model

Prague Approach Area

Baseline (Full ABSR)

Recognition Rate
92.1%
Error Rate:
0.60%

Delta

Recognition Rate: -3.5%
Error Rate: +0.48%

ABSR with Speaker-Independent AM

Recognition Rate
88.6%
Error Rate:
1.07%

Vienna Approach Area

Baseline (Full ABSR)

Recognition Rate
83.3%
Error Rate:
3.21%

Delta

Recognition Rate: -0.7%
Error Rate: +0.21%

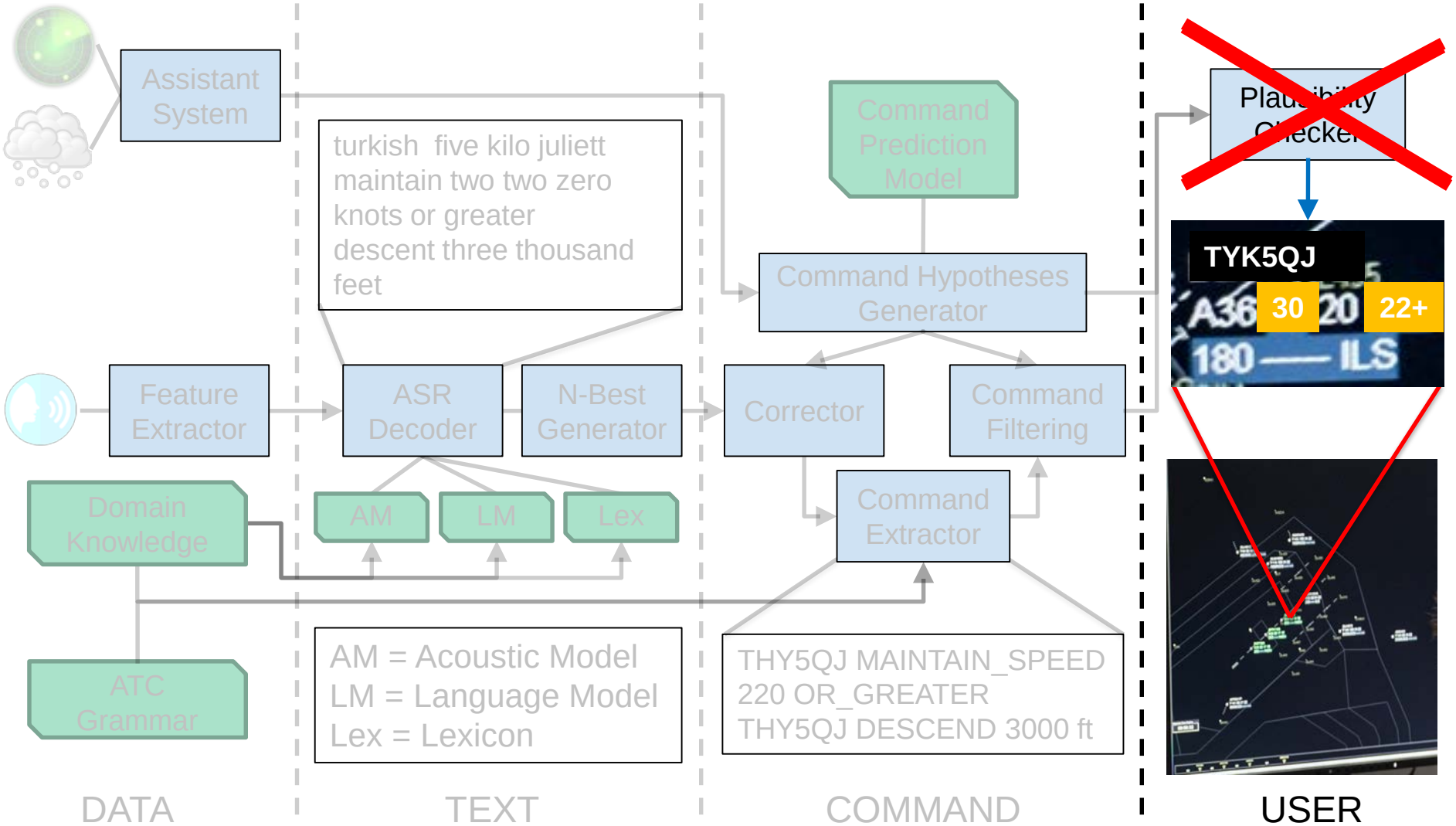
ABSR with Speaker-Independent AM

Recognition Rate
82.6%
Error Rate:
3.42%

DLR

MALORCA

Influence of Plausibility Checker



DLR

MALORCA

Influence of Plausibility Checker

Prague Approach Area

Baseline (Full ABSR)

Recognition Rate
92.1%
Error Rate:
0.60%

Delta
Recognition Rate: +1.9%
Error Rate: +1.23%

ABSR without Plausibility Checker

Recognition Rate
94.0%
Error Rate:
1.83%

Vienna Approach Area

Baseline (Full ABSR)

Recognition Rate
83.3%
Error Rate:
3.21%

Delta
Recognition Rate: +1.6%
Error Rate: +3.00%

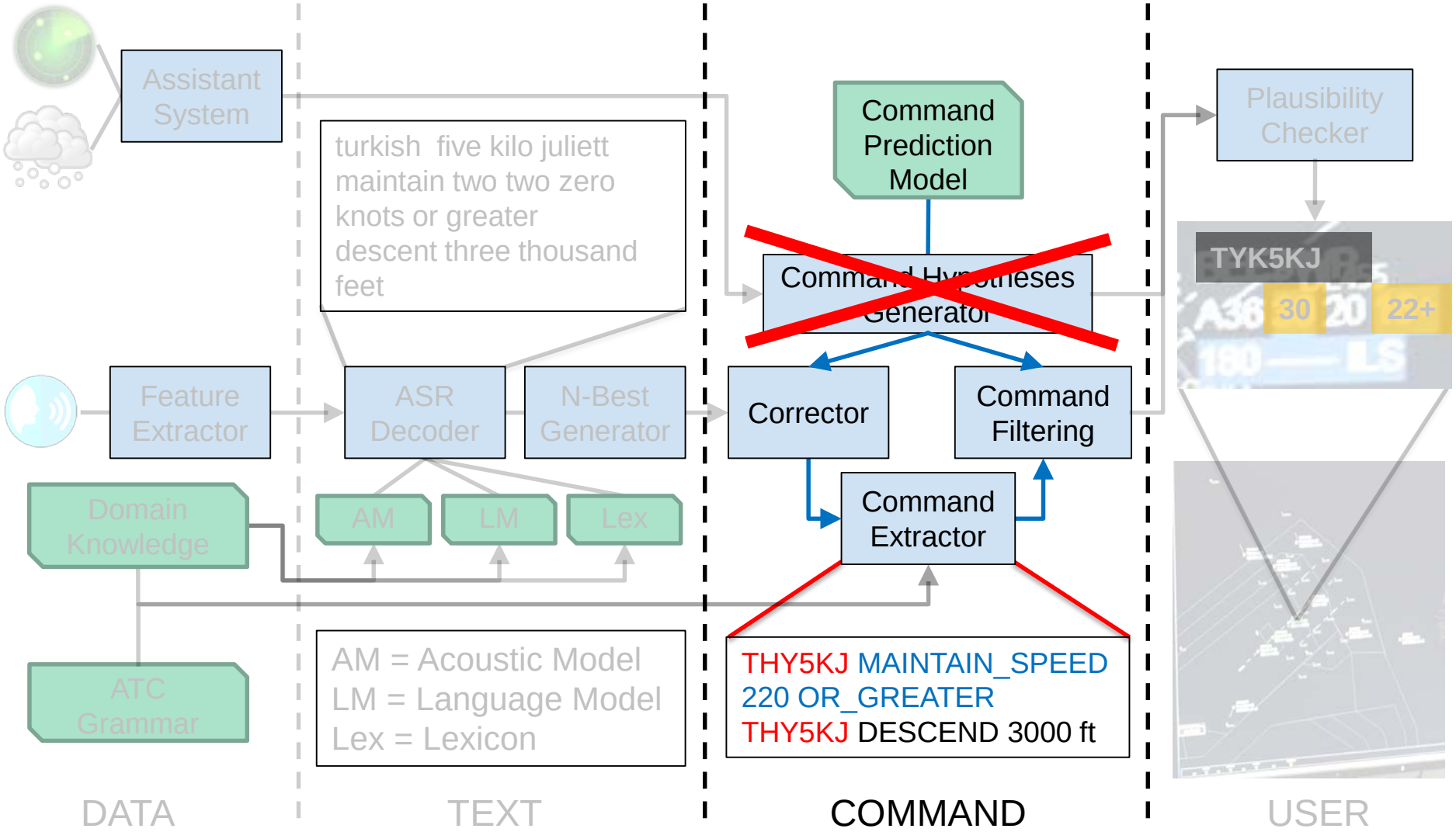
ABSR without Plausibility Checker

Recognition Rate
82.6%
Error Rate:
6.21%

DLR

MALORCA

Influence of Command Hypotheses Generator



DLR

MALORCA

Influence of Command Hypotheses Generator

Prague Approach Area

Baseline (Full ABSR)

Recognition Rate
92.1%
Error Rate:
0.60%

Delta
Recognition Rate: -4.5%
Error Rate: +6.07%

ABSR without Hypotheses Generator

Recognition Rate
87.5%
Error Rate:
6.66%

Vienna Approach Area

Baseline (Full ABSR)

Recognition Rate
83.3%
Error Rate:
3.21%

Delta
Recognition Rate: -11.8%
Error Rate: +11.78%

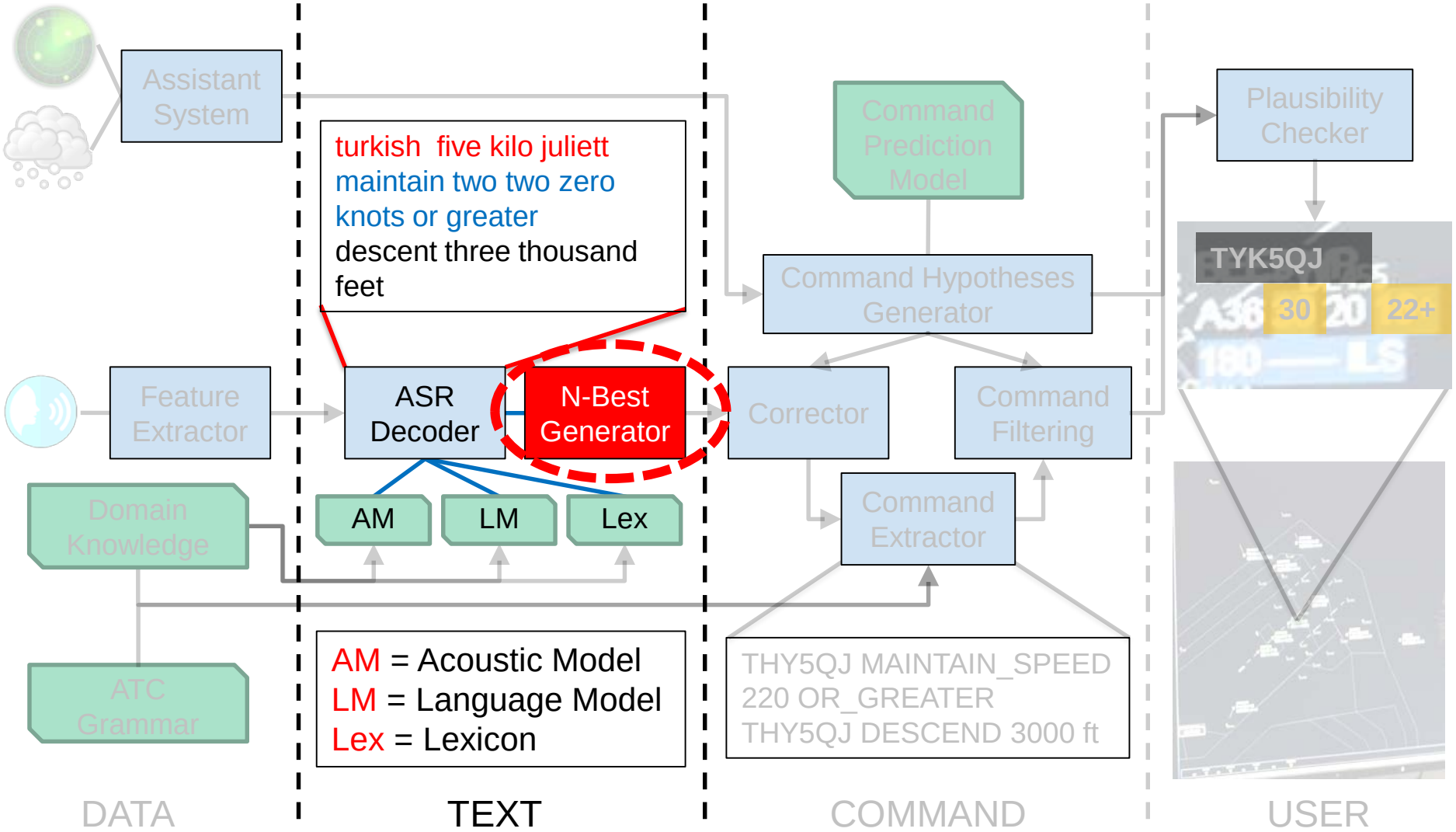
ABSR without Hypotheses Generator

Recognition Rate
71.5%
Error Rate:
15.00%

DLR

MALORCA

Influence of N-Best Generator



DLR

MALORCA

Influence of N-Best Generator

	Prague Approach Area	Vienna Approach Area
N-Best = 1	Recognition Rate: 91.4% Error Rate: 0.56%	Recognition Rate: 80.8% Error Rate: 2.85%
N-Best = 5 (Baseline)	Recognition Rate: 92.1% Error Rate: 0.60%	Recognition Rate: 83.3% Error Rate: 3.21%
N-Best = 10	Recognition Rate: 92.2% Error Rate: 0.57%	Recognition Rate: 83.7% Error Rate: 3.21%
N-Best = 20	Recognition Rate: 92.1% Error Rate: 0.73%	Recognition Rate: 83.7% Error Rate: 3.22%
N-Best = 50	Recognition Rate: 92.0% Error Rate: 0.62%	Recognition Rate: 83.9% Error Rate: 3.16%

DLR

Conclusion

- The use of **Machine Learning** for an ABSR-System is possible
- **ABSR** divided into different conceptual modules with
  - Reusable Building Blocks (No adaptation necessary)
  - Automatically learned Models (Acoustic-, Language-, Command Prediction Model)
  - Partially manual updates of Data elements
- Not all parts of the system have to be implemented (if decrease in performance is acceptable)
- Command Hypotheses Generator has the biggest influence on the overall performance

MALORCA

Building Blocks of Assistant Based Speech Recognition for Air Traffic Management Applications

Thank you for your attention!

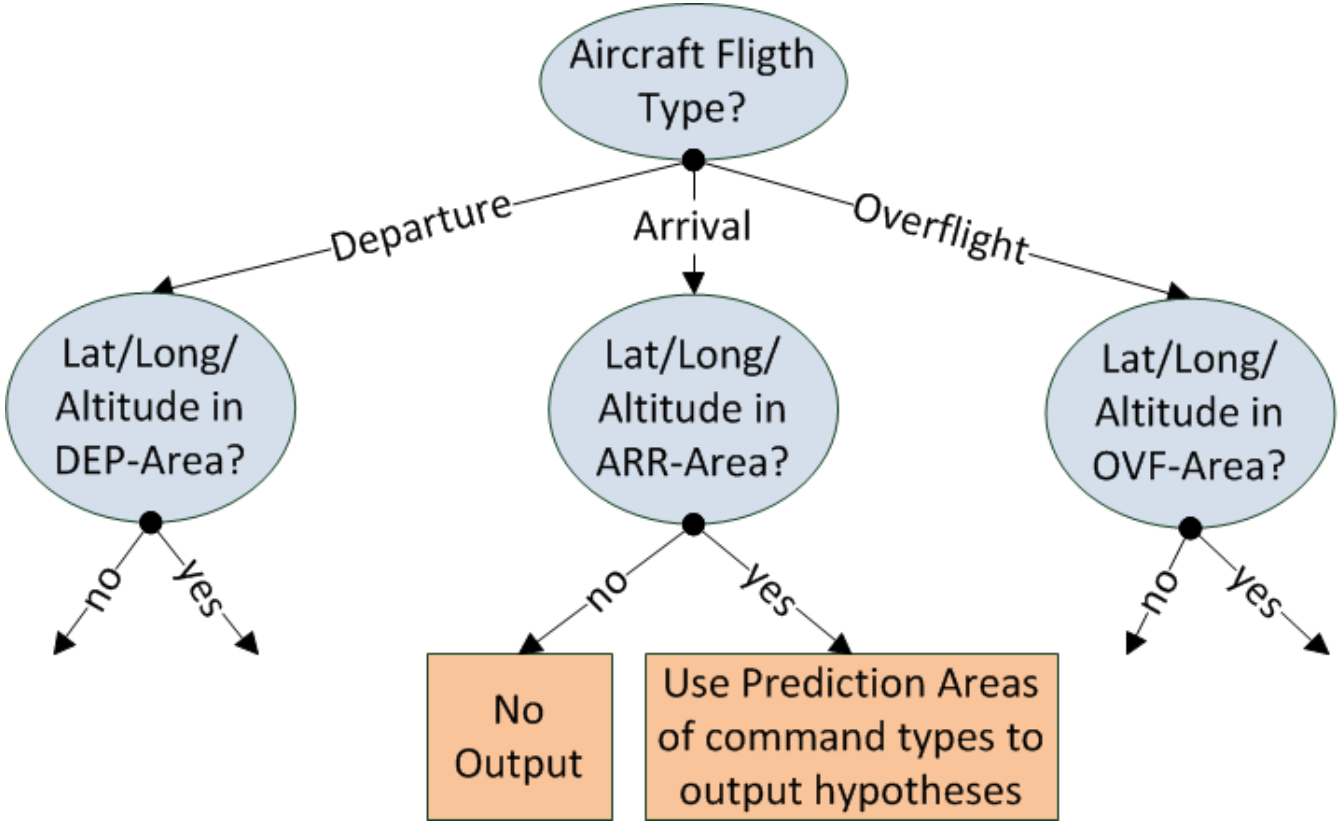
Matthias Kleinert, Hartmut Helmke
DLR, Institute of Flight Guidance
Controller Assistance

Matthias.Kleinert@dlr.de
Phone: +49 531 295-2567

Hartmut.Helmke@dlr.de
Phone: +49 531 295-2599

DLR

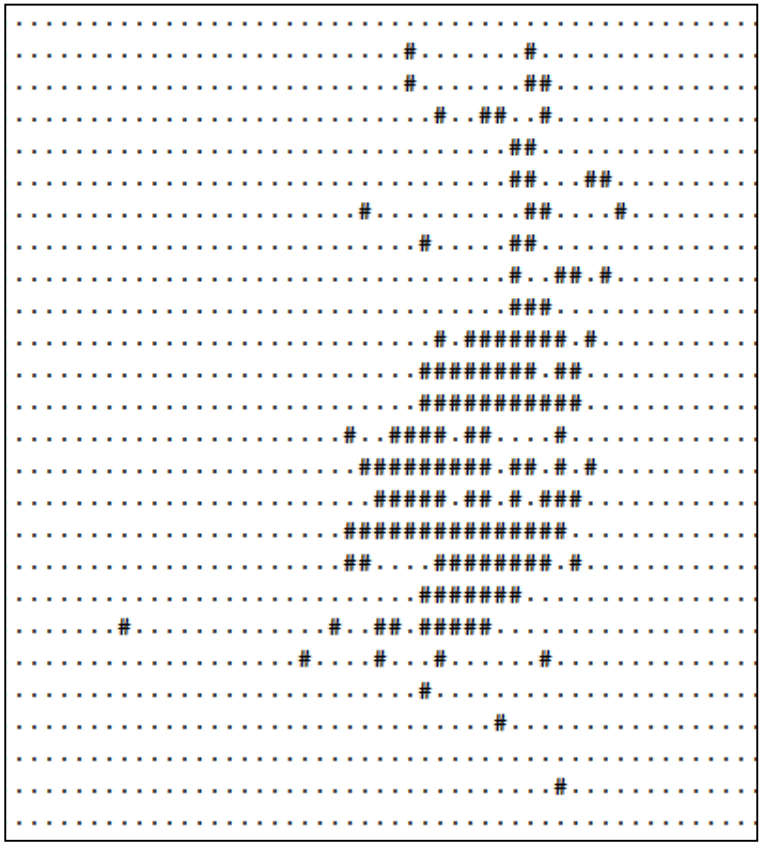
Commmand Prediction Model as Decision Tree



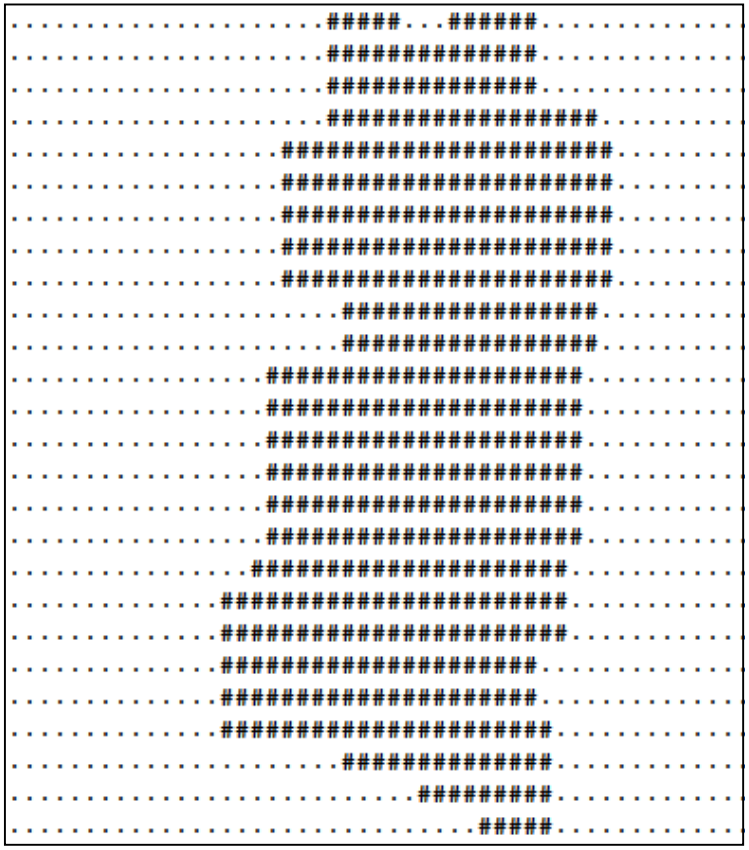
- Specific prediction areas for every command type (e.g. DESCEND, REDUCE, CLEARED ILS)

MALORCA

Expanding learned areas with different windows



Before Expansion



After Expansion

Prediction Area of CPM for Cleared ILS commands of Arrivals

DLR

MALORCA

Experimental Set-Up for Proof-of-Concept

- Real world data provided ANS CR and Austro Control
- Four different controller positions
- 18.7h training data for Vienna and 18.1h trainings data for Prague

Configuration	# Total Cmds	# Descend cmds	# ILS clearances
AEC (Prague)	11103	2184	351
PEC (Prague)	5365	920	458
BALAD (Vienna)	5929	1062	13
Feeder (Vienna)	6959	1100	245

Automatically transcribed (used for learning)

Approach Area	# Utterances	# given commands	# sessions
Prague	2582	4563	27
Vienna	2427	3556	19

Transcribed manually (test data)

DLR

User Acceptance (2)

Prague:				
Number of Commands	Number of ABSR Errors/Rejections	Rec Rate	corrected by controller	detected by controller
396	36	90.9%	31	36

Vienna:				
Number of Commands	Number of ABSR Errors/Rejections	Rec Rate	corrected by controller	detected by controller
610	80	86.9%	79	80

No safety issues were observed.

All misrecognitions were detected.

Better Recognition Rate of course would improve even more workload reduction.

MALORCA

Building Blocks of Assistant Based Speech Recognition for Air Traffic Management Applications

Matthias Kleinert, Hartmut Helmke, Heiko Ehr, Gerald Siol (DLR)
Christian Kern (Austro Control)
Dietrich Klakow, Mittul Singh (UdS)
Petr Motlicek (Idiap)

German Aerospace Center (DLR), Braunschweig

Saarland University (UdS), Saarbrücken

Idiap Research Institute, Martigny

Austro Control, Vienna

DLR