International Telecommunication Union

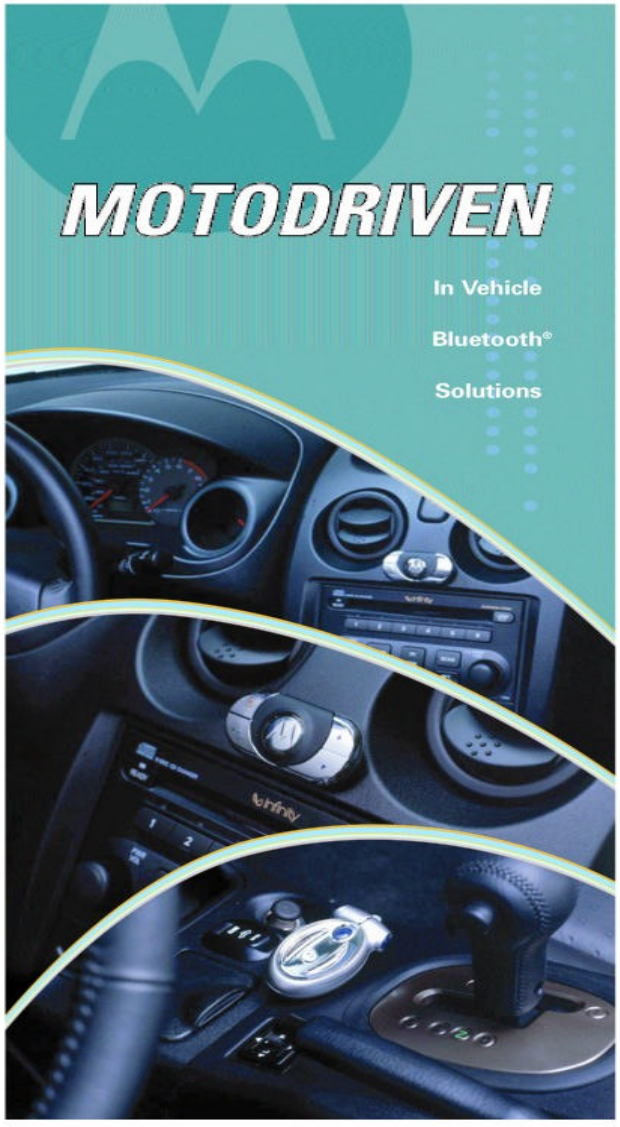
Perspectives on the use of Distributed Speech Recognition for in-car Telematics

David Pearce

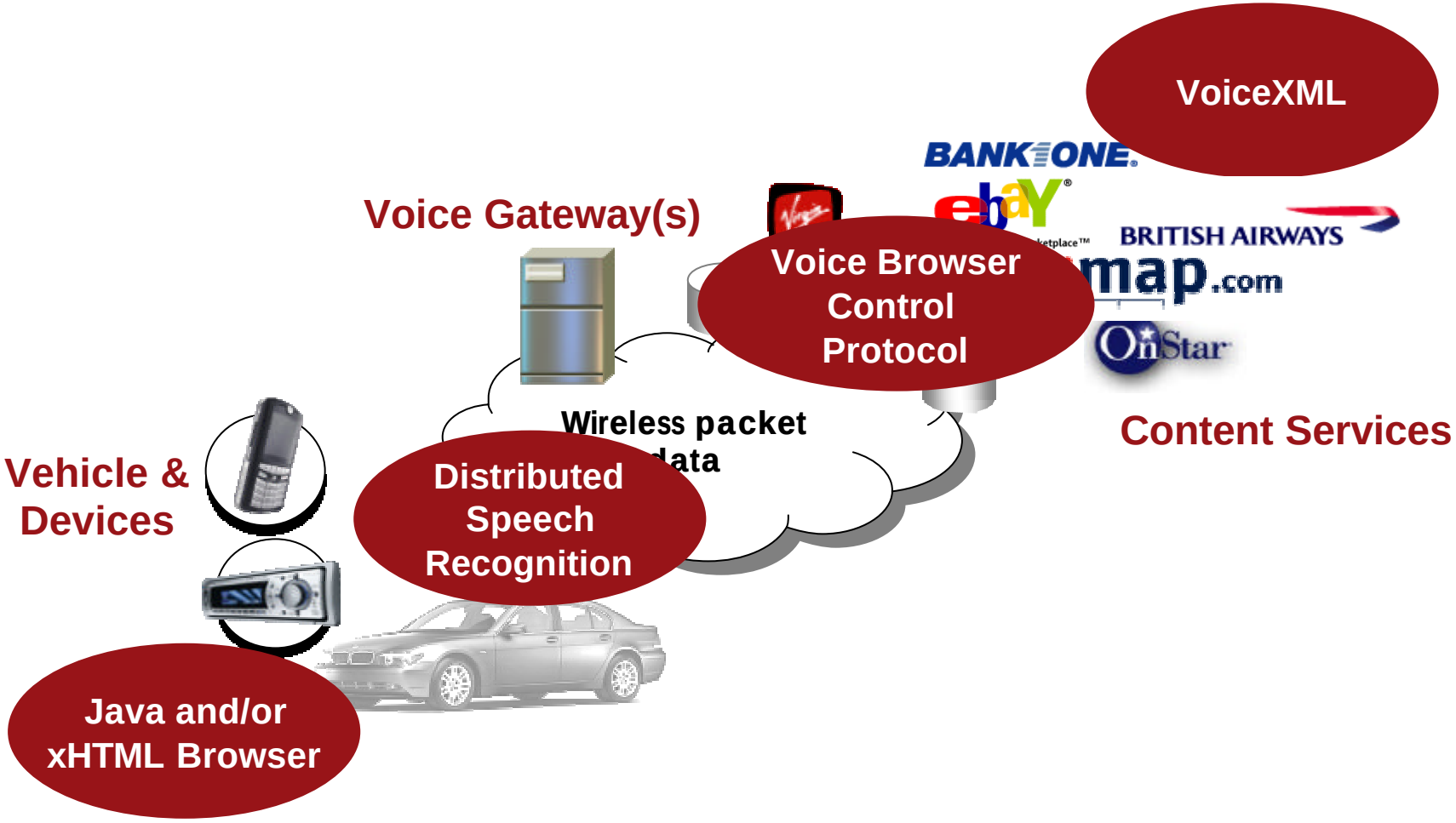
Motorola

"The Fully Networked Car, A Workshop on ICT in Vehicles"
ITU-T Geneva, 2-4 March 2005

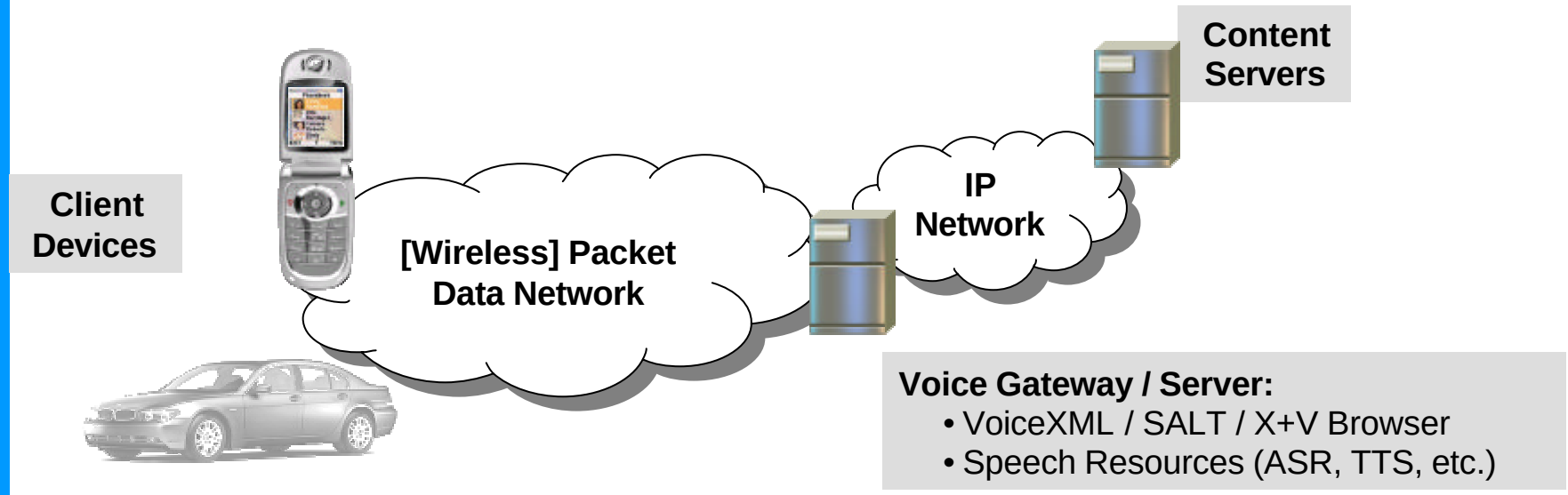
Embedded vs Connected



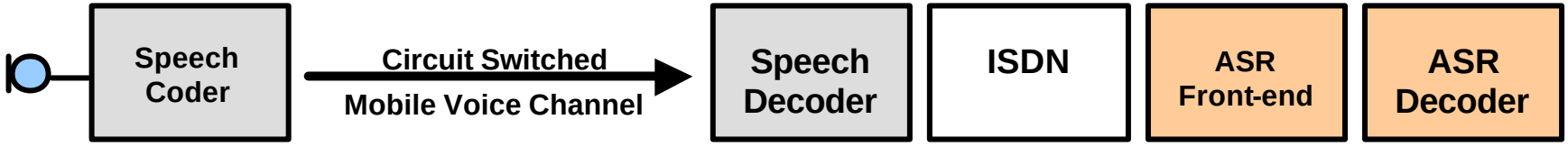
Key Elements of a Multimodal Architecture



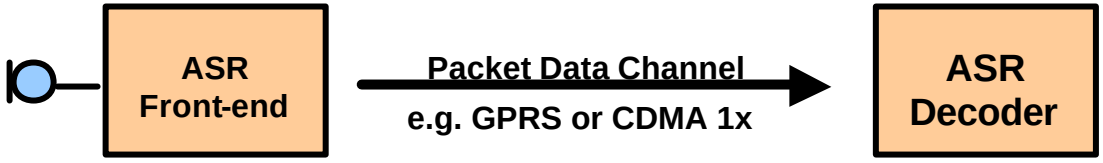
Distributed Speech Recognition



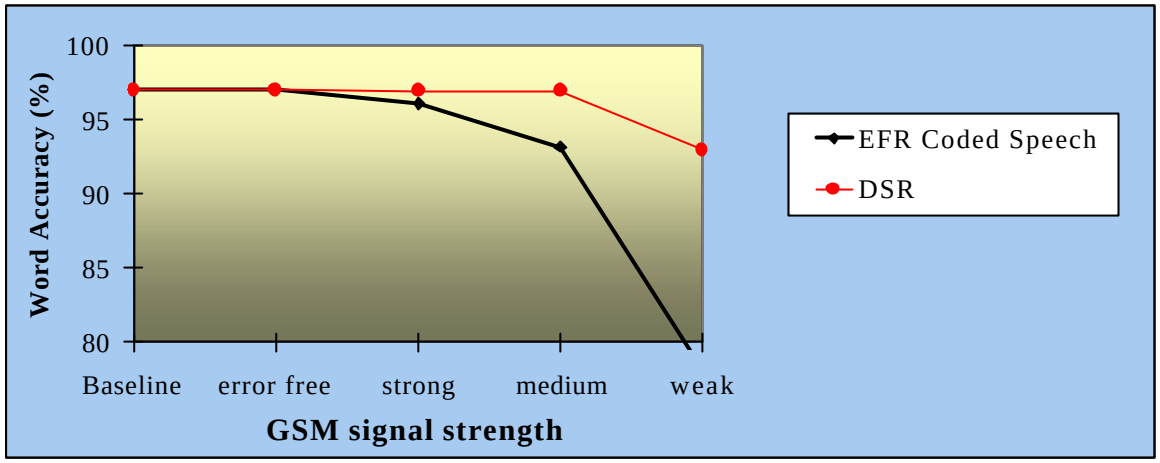
Conventional



DSR



Benefits of DSR



- **Improves performance over wireless channels**
  - **Minimises impact of codec & channel errors**
  - **Consistent performance over coverage area**
- **Improved performance in background noise**
  - **53% reduction in error rate**
- **Ease of integration of combined speech and data applications**
  - **Use packet data channel for both DSR and other data**

DSR Standards

ETSI

DSR Advanced front-end (Oct 02)
DSR Extended Advanced Front-end (Nov 03)

3GPP
A GLOBAL INITIATIVE

Speech Enabled Services
Fixed point DSR standard created
DSR selected as the recommended codec for SES
(Approved June 04)

IETF

RTP payload formats for DSR
Specifications standardised at rfc

3GPP2

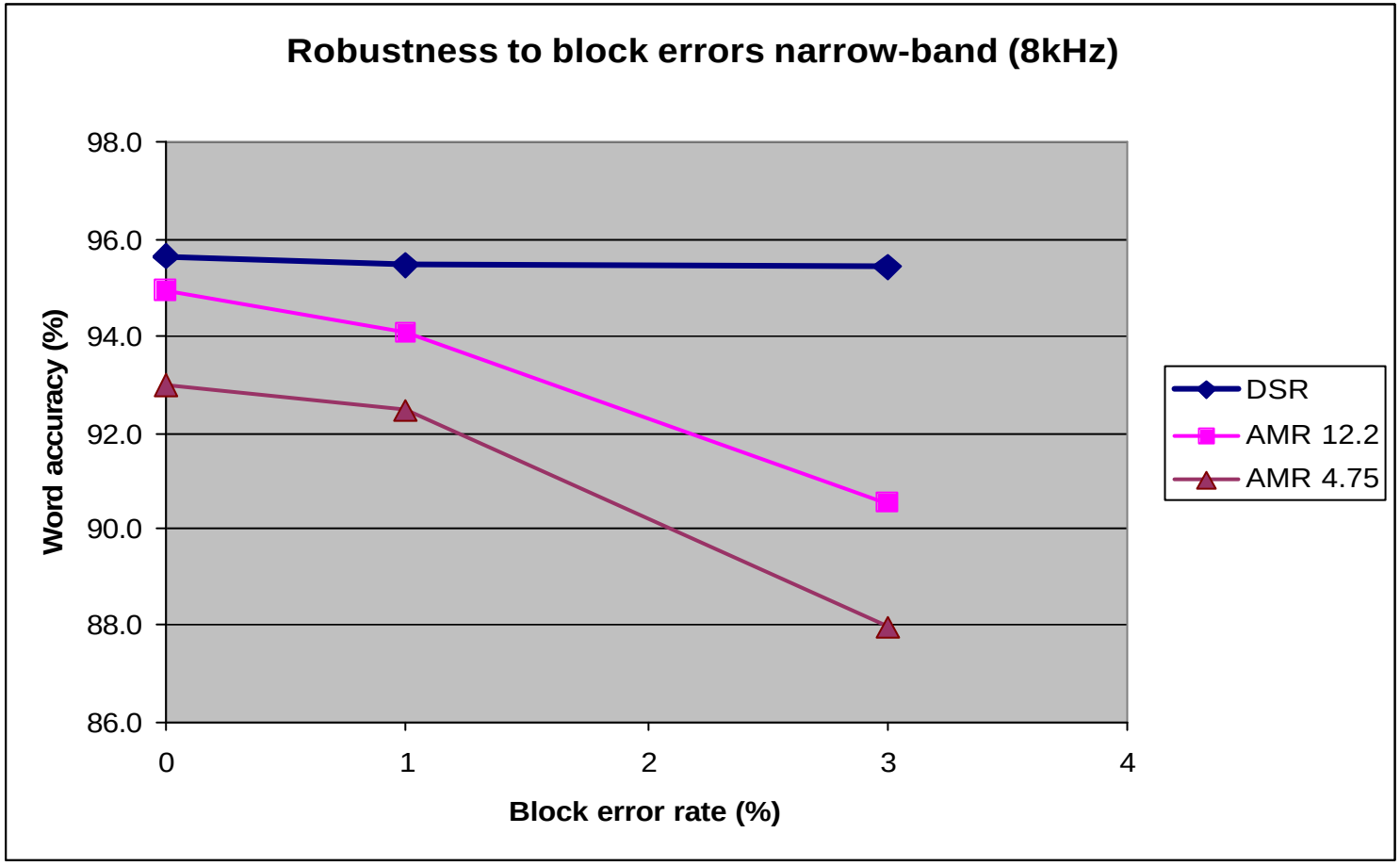
Speech Enabled Services
New Work Item (Approved Jan 05)

Results of ASR vendor evaluations in 3GPP

8 kHz	Number of db tested	AMR4.75 Average Absolute Performance	DSR Average Absolute Performance	Average Improvement
Digits	11	13.2	7.7	**39.9%**
Sub-word	5	9.1	6.5	**30.0%**
Tone confusability	1	3.6	3.1	**14.8%**
Channel errors	4	6.1	2.4	**52.8%**
Weighted Average				**36%**

- Extensive testing on 21 different speech databases
  - Covering different languages, tasks and environments
- Tests performed with IBM and Scansoft commercial recognisers
- Results for low data-rate comparison for packet data (< 8kbit/s)

Packet Switched Channel Errors



- **Aurora-3 Italian speech database**
- **GPRS network simulation for distribution of errors**

3GPP Feb 2004

Coded speech vs DSR

	DSR	AMR 4.75	Degradation
Well matched	96.5	94.4	-57%
Med mismatch	90.4	83.9	-68%
High mismatch	88.6	76.8	-104%
Average	**92.4**	**86.3**	**-73%**

	DSR	EVRC	Degradation
Well matched	96.5	90.6	-165%
Med mismatch	90.4	75.9	-151%
High mismatch	88.6	70.5	-160%
Average	**92.4**	**80.4**	**-159%**

Conclusions

- DSR provides substantial performance advantages for Voice driven Telematics services
- Standards for the DSR features and transport protocols for interoperability are complete
- Is there a need to incorporate in specific telematics standards?