Technical Specification Group Services and System Aspects
Meeting #13, Beijing, China, 24-27 September 2001

TSGS#13(01)0581

Work Item Description

Title

Speech Enabled Services Based on Distributed Speech Recognition (DSR)

1 3GPP Work Area

	Radio Access
X	Core Network
X	Services

2 Linked work items

End to End QoS (Concept and Architecture) for PS Domain (SA2)
Extended Transparent End-to-End Packet Switched Streaming Service (PSS-E) (SA4)
Packet-switched Conversational multimedia Applications (SA4)
IMS (SA1)

3 Justification

Forecasts show that speech-driven services will play an important role on the 3G market. People want the ability to access information while on the move and the small portable mobile devices that will be used to access this information need improved user interfaces using speech input. At present, however, the complexity of medium and large vocabulary speech recognition systems are beyond the memory and computational resources of such devices.

Distributed Speech Recognition (DSR) overcomes these problems, and it will provide 3G users with a high performance distributed speech interface to server-based automatic information and transactional services.

The types of services include those that are voice only, for example, automatic speech access to information. In the future, a new range of multi-modal applications is also envisaged incorporating different modes of input (e.g. speech, keyboard, pen) and speech and visual output.

4 Objective

4.1 To enable all these benefits in a wide market, such as 3G, containing a variety of players including terminal manufacturers, operators, 3rd Party Service Providers and recognition vendors, a standard for the FE is needed to ensure compatibility between the terminal and the remote recogniser. The first standard for a DSR front-end and compression was published by ETSI in Feb 2000.

4.2 In addition to the DSR front-end, a standard DSR protocol stack is needed to support end-to-end interoperability. ETSI STQ Aurora has also been developing proposals for these transport protocols that will be standardized by the IETF. DSR applications will be based on the IETF packet protocols using RTP (Real Time Protocol), SDP (Session Description Protocol) and SIP (Session Initiation Protocol).

3GPP will standardise the minimum to allow inter-operability.

5 Service Aspects

The WI will define the necessary components for speech enabled services based on Distributed Speech Recognition (DSR), for example automatic speech access to information.
This WI will identify the necessary changes and additions required in the current SA1 specifications.

6 MMI-Aspects

Man Machine Interface aspects have to be considered but not standardised.

7 Charging Aspects

Charging aspects have to be considered. Same as IMS charging.

8 Security Aspects

Security aspects have to be considered. Same as IMS.

9 Impacts

Affects:	USIM	ME	AN	CN	Others
Yes		x			
No	x				
Don't know				X	

10 Expected Output and Time scale (to be updated at each plenary)

New specifications						
Spec No.	Title	Prime rsp. WG	2ndary rsp. WG(s)	Presented for information at plenary#	Approved at plenary#	Comments
TS 22.xxx	DSR stage 1.			SA#14 Kyoto	SA#15	

Affected existing specifications				
Spec No.	CR	Subject	Approved at plenary#	Comments
TR 22.941		Inclusion of DSR in IMS Framework Document	Done	
TS 23.xxx		Inclusion of DSR in stage 2.	tbd	May be a stand alone stage 2.
TS 23.207		Inclusion of DSR in QoS spec.	tbd	
TS 24.xxx?		SDP protocols extension to include DSR	tbd	Awaiting guidance from CN.

Note: ETSI ES 201 108: Existing ETSI specification “Speech processing, Transmission and Quality Aspects (STQ); Distributed Speech Recognition; Front-end feature extraction algorithm; Compression algorithms” will be referenced by appropiate 3GPP TS(s).

11 Work item raporteurs

D Williams, QUALCOMM, Inc. Dwilliams@qualcomm.com

12 Work item leadership

TSG SA WG 1

13 Supporting Companies

Alcatel, Motorola, Qualcomm, France Telecom, Texas Instruments, Vodafone, Mannesmann, Omnitel, IBM, Sony.

14 Classification of the WI (if known)

	Feature (go to 14a)
x	Building Block (go to 14b)
	Work Task (go to 14c)

14a The WI is a Feature: List of building blocks under this feature

(list of Work Items identified as building blocks)

14b The WI is a Building Block: parent Feature

Speech Recognition and Speech Enabled Services

14c The WI is a Work Task: parent Building Block

(one Work Item identified as a building block)