FP6–004171 HEARCOM
Hearing in the Communication Society

INTEGRATED PROJECT
Information Society Technologies

D-9-5: Definition of PCS platform and assistive communication services

Contractual Date of Delivery: 31 August 2006(+45 days)
Actual Date of Submission: 16 October 2006
Editor: FR-TCF
Sub-Project/Work-Package: SP4/WP9
Version: 2.0
Total number of pages: 73

Dissemination Level

<table>
<thead>
<tr>
<th>PU</th>
<th>Public</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PP</td>
<td>Restricted to other program participants (including the Commission Services)</td>
<td></td>
</tr>
<tr>
<td>RE</td>
<td>Restricted to a group specified by the consortium (including the Commission Services)</td>
<td></td>
</tr>
<tr>
<td>CO</td>
<td>Confidential, only for members of the consortium (including the Commission Services)</td>
<td></td>
</tr>
</tbody>
</table>

Project co-funded by the European Commission within the Sixth Framework Program (2002-2006)
This information is confidential and may be used only for information purposes by Community Institutions to whom the Commission has supplied it
## Deliverable D-9-5

### VERSION DETAILS

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.0</td>
<td>2006</td>
<td>Final</td>
</tr>
</tbody>
</table>

### CONTRIBUTOR(S) to DELIVERABLE

<table>
<thead>
<tr>
<th>Partner</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>FR-TCF</td>
<td>Gwénaël Guilmin</td>
</tr>
<tr>
<td>ES-MOV</td>
<td>Carolina Rincón, Javier Sainz</td>
</tr>
<tr>
<td>GR-ILSP</td>
<td>Akis Athanaselis</td>
</tr>
<tr>
<td>DE-OFFIS</td>
<td>Frank Poppen</td>
</tr>
<tr>
<td>UK-RNID</td>
<td>Amy Izycky</td>
</tr>
<tr>
<td>NL-VUMC</td>
<td>Marcel Vlaming, Adriana Zekveld</td>
</tr>
</tbody>
</table>

### DOCUMENT HISTORY

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Responsible</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>19/6/06</td>
<td>Gwenaël Guilmin</td>
<td>Initial definition of chapters</td>
</tr>
<tr>
<td>0.2</td>
<td>14/7/06</td>
<td>Carolina Rincón</td>
<td>Contribution MOV</td>
</tr>
<tr>
<td>0.3</td>
<td>17/7/06</td>
<td>Gwenaël Guilmin</td>
<td>Contribution TCF</td>
</tr>
<tr>
<td>0.4</td>
<td>18/7/06</td>
<td>Frank Poppen</td>
<td>Contribution OFFIS</td>
</tr>
<tr>
<td>0.5</td>
<td>26/7/06</td>
<td>Marcel Vlaming</td>
<td>Contribution VUMC</td>
</tr>
<tr>
<td>0.6</td>
<td>28/7/06</td>
<td>Akis Athanaselis</td>
<td>Contribution ILSP</td>
</tr>
<tr>
<td>0.7</td>
<td>30/8/06</td>
<td>Amy Izycky</td>
<td>Contribution RNID</td>
</tr>
<tr>
<td>0.9</td>
<td>31/8/06</td>
<td>Javier Sainz</td>
<td>Updated with OFFIS contribution</td>
</tr>
<tr>
<td>0.10</td>
<td>19/9/06</td>
<td>Gwenaël Guilmin</td>
<td>Updated with TCF contribution</td>
</tr>
<tr>
<td>0.12</td>
<td>20/9/06</td>
<td>Marcel Vlaming</td>
<td>Modif. architecture of the doc</td>
</tr>
<tr>
<td>0.13</td>
<td>22/9/06</td>
<td>Javier Sainz</td>
<td>Additional info from MOV</td>
</tr>
<tr>
<td>0.14</td>
<td>24/9/06</td>
<td>Gwenaël Guilmin</td>
<td>Prepared for review</td>
</tr>
<tr>
<td>1.0</td>
<td>11/10/06</td>
<td>Gwenaël Guilmin</td>
<td>Updated according to J. Koopman comments</td>
</tr>
<tr>
<td>2.0</td>
<td>14 October</td>
<td>M. Vlaming</td>
<td>check</td>
</tr>
</tbody>
</table>

### DELIVERABLE REVIEW

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Reviewed by</th>
<th>Conclusion*</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>15/9/2006</td>
<td>T. Hout gast</td>
<td>Rework/restructure</td>
</tr>
<tr>
<td>1.0</td>
<td>9/10/06</td>
<td>J. Koopman</td>
<td>Modify / accept</td>
</tr>
</tbody>
</table>

* e.g. Accept, Develop, Modify, Rework, Update
## Table of Contents

1. Preamble .................................................................................. 7
2. Executive Summary ................................................................... 7
3. Introduction .............................................................................. 9
4. PCS platform ............................................................................11
   4.1 Description .........................................................................11
   4.2 PCS Prototype System .........................................................14
   4.3 Requirements for PCS ..........................................................15
   4.4 Overview on available mobile PDA’s .......................................22
   4.5 Description of selected PDA for demonstration .........................24
   4.6 Communication services .......................................................30
   4.7 Hardware extensions ...........................................................31
5. Assistive communication services ................................................33
   5.1 Public Announcement and Address Systems .........................33
   5.2 Speech recognition system (client-server) ...............................38
   5.3 Personal Hearing System ......................................................44
   5.4 Assistive information display ...............................................48
   5.5 Service Interactions .............................................................54
6. Software Components ...............................................................57
   6.1 Extension and adaptations for Software .................................57
7. Dissemination and exploitation ...................................................59
8. Ethical Issues ...........................................................................59
9. Conclusions ..............................................................................60
10. References .............................................................................61

Appendix 1: International Phonetic Alphabet .................................62
Appendix 2: SAMPA computer readable phonetic alphabet ..........63
List of Figures

Figure 1 Overview of PCS platform. The Personal Communication Link will be based on a Wireless Body Area Network (WBAN).

Figure 2 Overview of the prototype PCS platform in which the PHS is an external device and in which the PCL/WBAN is included as a simulation by wired interfaces and Blue-tooth.

Figure 3 Market share of PDA Operating Systems.

Figure 4 POCKET PC IPAQ HP6340.

Figure 5 OMAP processor.

Figure 6 Innovator kit platform.

Figure 7 Architecture for Public Announcement System.

Figure 8 Public announcement/address system.

Figure 9 Demo Proposal PA system.

Figure 10. Left: ASR Client on PDA; Right: After being connected to the server.

Figure 11 Display of ASR server output.

Figure 12 HEARCOM Client Server Elements.

Figure 13 ASR Server Functions.

Figure 14 ASR and TTS Functions.

Figure 15 Two scenarios of PHS-usage.

Figure 16 Scenario one prototype.

Figure 17 Scenario two prototype.

Figure 18 Rolling Text Elements by Scroll.

Figure 19 Rolling text Human Machine Interface (example).

Figure 20 Display Options.

Figure 21 Rolling Text Options.
Figure 22 Rolling Text Options ................................................................. 53
Figure 23 GPRS architecture ................................................................. 68
Figure 24 SDIO cards with external functions ......................................... 70
Figure 25 Public Announcement Block Diagram ....................................... 71

List of Tables

Table 1 Delay requirements ..................................................................... 16
Table 2 Signal quality requirements ......................................................... 17
Table 3 PCS Physical requirements ......................................................... 18
Table 4 WAN Communication link scenario ............................................ 18
Table 5 PAN Communication link scenario ............................................. 19
Table 6 PCL Communication link scenario ............................................. 19
Table 7 Hardware requirements for PCS platform .................................. 21
Table 8 Mobile PDA devices overview (from manufacturer’s web sites) ....... 24
Table 9 HP iPAQ Pocket PC HP6340 specifications .................................. 26
Table 10 Functionalities of ASR System ................................................... 40
Table 11 HMI Rolling text display functionalities ..................................... 54
Table 12 Table of Software Components ............................................... 57
Abbreviations

3G  See UMTS
ASR  Automatic Speech Recognition
BAN  Body Area Network
BT  Bluetooth
DSP  Digital Signal Processor
CF  Compact Flash
EDGE  Enhanced Data rates for Global Evolution
GPRS  General Packet Radio Service
GPP  General Purpose Processor
GPS  Global Position System
GSM  Global System for Mobile Communications
HA  Hearing Aid
HMI  Human Machine Interface
I/O  Input/Output
LAN  Local Area Network
MFLOPS  Million Floating Point Operations per Second
MIPS  Millions Instruction Per Second
MMI  Man Machine Interface
OS  Operating System
PAN  Personal Area Network
PCL  Personal Communication Link
PCS  Personal Communication System
PDA  Personal Digital Assistant
PHS  Personal Hearing System
RF  Radio Frequency
RIM  Research in Motion's
SD card  Secure Digital card
SDIO  Secure Digital card with IO interface
TCP  Transmission Control Protocol
UDP  User Datagram Protocol
UMTS  Universal Mobile Telephone Service
USB  Universal Serial Bus
WAN  Wide Area Network
WBAN  Wireless Body Area Network
WPA  Wireless Public Address
WiFi  Wireless Fidelity, IEEE standard 802.11b

Acknowledgement

Supported by grants from the European Union FP6, Project 004171 HEARCOM. The information in this document is provided as is and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability.
1 Preamble

This deliverable is an overview document on the Personal Communication System (PCS) and reports on tasks 9.2 and 9.3 of the WP9 workplan:

- Based on the requirements of (D-9-1) and the Personal Communication Link (PCL) of WP8, a prototype Personal Communication System (PCS) platform (commercial PDA) will be selected. Adaptations for demonstrating the assistive services will be specified.

- Communication services for the PCS will be selected and defined. The services will focus on the interaction and connection of the PCS to mobile phones, public-address and the client server speech text conversion system.

This deliverable will relate to the following deliverables:

- D-8-4 for an overview on the technical solutions of the Personal Communication Link (PCL) that will provide a wireless connection between PCS and hearing aids.

- D-9-3 and D-9-4 for the specification, design and implementation of the Personal Hearing System (PHS). The PHS is intended to be integrated to the PCS, but for demonstration will be implemented onto a separate personal notebook. The PHS will be used in SP3 as a platform for the testing and evaluation of hearing algorithms.

- D-9-7 and D-9-8 for the definition and specification of a client-server ASR application that will be included onto the PCS.

- D-9-6 for the research and development of an advanced text display method for use on a PCS. The research concerns the integration of audio and display information to assist communication.

In follow up deliverables the detailed specifications and design details will be provided.

2 Executive Summary

The Personal Communication System (PCS) is an assistive platform for hearing impaired persons that assists on communication tasks. On this platform a collection of assistive services will be based.

This deliverable will cover the selection and the definition of the PCS platform and the description of new defined assistive communication services. These services will focus on the connection of the PCS to wireless public-address services, telephone usage using a client server speech to
text conversion system and the integration of a personal hearing system (PHS) for enhanced audio and speech processing. The PCS will provide connection to hearing devices (hearing aids or headsets) using a personal communication link as defined in WP8 (D-8-4).

This deliverable first gives a description of the personal communication system and the definition of prototype configuration that will be used to provide a demonstrator on the functionalities of the PCS. Requirements of the PCS are discussed and a selection is made of a PDA on which the prototype PCS will be based.

In a next section the assistive services are defined and described. The assistive services for the prototype PCS include:

- A wireless public announcement system (WPA) that provides public wireless announcements for a certain area that can be retrieved by listening or be text presentation.

- An automatic speech recognition system (ASR) that assists on phone conversations, based on a client server application on the Internet. This service will include a feedback module for structure and control by which a conversation will be structured to the level on which speech recognition will be possible at a reliable level.

- A personal hearing system (PHS) that will provide personalized sound and speech processing for hearing impaired persons. This system can be used in addition to hearing aids. This system (excluding other PCS functionalities) will be applied in WP5 for development of hearing algorithms; In WP7 user tests will be performed using the PHS.

- An advanced system for the combined presentation of audio and textual (bimodal) speech information that will enhance speech communication under adverse acoustical conditions (background noise, reverberation etc). This bimodal system will increase the performance of the wireless announcements service and the speech recognition service.

The deliverable concludes with an overall specification of required software modules. The specification of the modules itself will be handled in later deliverables for each specific service.
3 Introduction

The objectives of subproject 4 are the development and the evaluation of technical devices, compatible with mainstream developments, to support communication in adverse conditions. To do that, this subproject aims firstly at the definition of a wireless link that can be used for hearing devices and that will be compatible with mainstream developments, and secondly at the investigation and prototyping of assistive applications including the definition of a common platform.

The selected assistive applications will support speech understanding by rolling text display, speech text conversion and speech processing using hearing devices. These applications will be implemented on a proposed technical handheld solution or Personal Communication System (PCS) that connects to hearing devices. This PCS is the main device in the center of this concept and will be based on a PDA terminal with multiple wireless functionalities and communication links to external devices and services.

The PCS will provide the interconnection between hearing aids based on a Wireless Body Area Network (WBAN) and Wide Area Network (WAN) devices, in order to offer enhanced services for hearing impaired people.

This document is divided into two main parts:

1. Chapter 4 defines the PCS platform, the selected PDA.
2. Chapter 5 describes the assistive communication services

The goal of this deliverable is to develop the idea’s on the PCS concept, and defining the functionalities of a demonstrator PCS. Details on the selected PCS communication services will be described in related reports.

The planned architecture of the HEARCOM communication system

The main function of the PCS is to connect the hearing aids with services and also to perform assistive signal processing for the hearing aids. It is directly connected to the mobile telephone network, and to the fixed-line telephone network, and to radio and television through wireless networks
such as WiFi or Bluetooth, selecting whatever provides the desired service. The PCS can also receive information for announcements (e.g. in a train station or airport). For additional support, the PCS can display textual information simultaneously with auditive speech to obtain a combined reinforcing effect on intelligibility.
4 PCS platform

4.1 Description

The Personal Communication System (PCS) is a hand held device that acts as a communication concentrator between the Personal Communication Link (PCL) at one side, and wireless communication inputs from the outside world at the other side:

- The PCL is defined as a short range wireless audio link between the PCS and hearing devices. The short range link will be based on a Wireless Body Area Network (WBAN), see D-8-4.

- Communication from the outside world which ranges from audio input, telephone input, alarms, assistive services, announcements, entertainment etc. The PCS acts as a bridging function that will assist the hearing impaired person in communication.

Figure 1 provides an overview on the services of the PCS system and its interfaces to hearing devices and communication inputs.

The PCS system has the potential to support many assistive and general services and applications. Specific applications for the HEARCOM project include the Personal Hearing System (PHS), the Automatic Speech Recognition System (ASR), a wireless public address (WPA) service and the bimodal presentation of information (text and speech).

The PHS is an optional system for running hearing aid algorithms that will enhance speech amplification for use by hearing impaired persons. This approach gives the possibility of using more powerful algorithms than can currently run on hearing devices. In addition the PHS offers a platform for development and testing as will be used in SP3.

The ASR service will be a client-server system for the support of phone conversations by providing a speech to text conversions that will be displayed on the PCS. Alternatively this can also be applied for acoustical inputs from microphone or Public Address systems. The system will send this audio signal to a server system for processing, and return the ASR output to the screen. The display on the screen will be via a rolling text system. As an alternative to the server based ASR system, it should be possible to route the signal to a service centre where speech to text operators could produce the output. This opens up the potential for future remote applications of speech to text services, and also provides a more reliable (if more expensive) solution than ASR technology that users may desire for important conversations.

The wireless public address system (WPA) is a new proposed public service that provides wireless access to announcements information in
public transport (train, airports etc), churches, schools, theatres etc. The PCS will receive location specific information using the local public Wifi connections and will display in textual mode or will replay in enhanced audio format at the hearing devices. Information will also be stored and can be accessed and replayed on demand.

In addition the PCS as most modern communication devices have built in capability to receive email, Internet pages, and mobile phone functions. In option the PCS device may be equipped by tactile or visual alerting options.

**Figure 1 Overview of PCS platform. The Personal Communication Link will be based on a Wireless Body Area Network (WBAN).**
A number of characteristics required to perform the functions and services for the HEARCOM project will be defined in this document.

4.1.1 PCS Application outline

The goal of the PCS is to improve the quality of communication that hard of hearing people experience in their personal and professional environments, as well in public places. The final target is that this device is, for any person, as simple to carry as a smart mobile phone or PDA.

When the person uses the PCS to link with their Hearing Aids (HA’s) or headset, the system has to adapt the signal to the specific requirements of the person. This will enable the user to obtain a better quality of sound when listening to the music that plays from their music system, watching TV, or when having a conversation with other people. This system will also help people to detect the alarms that surround us in daily life, from the telephone ring and the door bell to those programmed in new smart home appliances and also alarms that may sound when in a public place. During public events such as theatre performances, religious services, conferences or courses, the system helps the end-user to enjoy these events by adapting the acoustic signal to his or her needs. During this situation, the end-user can also extract this device from the pocket to check if there is further information available referring to the act, for instance sent by the organizer through the wireless network (general information, concept indication, text transcription by speech-text converter integrated in the system which allows reading it in the format of a rolling display), and could extend to many places (stations, airports, and subway trains) for public information announcements or personalized alerts.

The availability of text information (which can also be spoken out by the device) is very important also for blind people, e.g. in the airports that have suppressed the voice announcements in order to adopt the newly trend of becoming “silent airports”, as well for foreign speaking people.

Given all these applications and scenarios the PCS has to have the capacity of connecting to several wireless networks, which would then allow access to the enormous amount of information that is available (analog or digital radio stations, Internet through GPRS, UMTS or WiFi, information based on location with GPS tools, etc).
4.2 PCS Prototype System

Figure 2 Overview of the prototype PCS platform in which the PHS is an external device and in which the PCL/WBAN is included as a simulation by wired interfaces and Blue-tooth.

With the current state of technology and product development it is not possible as yet to integrate all intended functionalities in one complete PCS system. Within the scope of SP4, 2 important elements are missing for realizing the PCS:

1. PCL/WBAN radio components that are integrated within a PCS (and integrated in hearing devices; which are not on the market as yet). In WP8 the PCL/WBAN concept is investigated. It is planned that first WBAN components will be developed by a selected provider and extended with audio channels coded for hearing aid use.

2. Processing power that is sufficient to allow the practical implementation of the PHS into the PCS. In principal it would be possible to make a dedicated implementation for the PHS using a simple and fixed signal processing that can be run on a PDA.
processor or integrated DSP processor. This is not the intention as this will not allow for advanced and flexible signal processing schemes for hearing devices as will be developed and applied in SP3.

The Demonstrator PCS will therefore be modified as follows as shown in figure 2:

1. Simulation of PCL/WBAN by:
   a. Using wired interfaces to the hearing aids, or alternatively
   b. using Bluetooth headsets that simulate hearing aids, (and otherwise used for people without hearing impairment).

2. Application of a separate PHS device that provides ample processing power and will allow the flexible implementation of new and advanced hearing algorithms (as developed in SP3-WP5). By this the PHS will be applied as a wearable test device for use in SP3-WP7.

The separate PHS device will consist of a small notebook that will be connected by standard (IP) links to the PCS.

The other PCS services will be implemented on a PDA device. The selection of the PDA will be described in next sections. This PDA will be adapted to demonstrate the PCS functionality and to prove the capability to implement such services in a handheld device.

### 4.3 Requirements for PCS

The PCS requirements will be summarized for user requirements, system requirements and service requirements.

#### 4.3.1 User requirements

In D-9-1 user requirements for a PCS were analyzed by user and professional questionnaires in 3 countries: UK, NL and GR.

The most important scenario’s for usage of the PCS concept were concluded as:

1. At transport terminals for announcements,
2. Inside public transport vehicles for announcements,
3. For telephone conversations,
4. At home for TV and radio,

5. In lectures or classes.

The willingness to carry an additional portable device that would enhance hearing and assist on communication was also surveyed upon. In general 94% (UK), 81% (NL) and 100% (GR, but small group) are willing to carry an additional device. In general it was concluded that across all three population samples, the majority of users would be willing to carry an additional device if it was small enough to be held in the palm of their hand and integrated with a mobile phone.

4.3.2 Functional requirements

4.3.2.1 Delay requirements

In the assistive hearing impaired domain, the delay between the original acoustic signal (which reaches directly the user) and the processed signal (which passed through the PCS before to reach the user) has to be reduced to the minimum to avoid any discomfort and annoyance.

In the hearing-aid context two types of delays are distinguished: processing delay and relative delay.

The processing delay is introduced in the signal processing chain between the microphone and the loudspeaker of the hearing aids, including transmission over the PCL that connects hearing aids and PHS. The tolerable processing delay depends on which aspect is tested (annoyance, echo, conversational problems, synchronization with visual impression, type of hearing aid, level/type of hearing loss,..). Anyway the overall processing delay can not be fixed and it is a trade-off between the annoyance due to the delay and what can be achieved with the signal processing. The maximum processing total delay allowed is set to 15 ms for the HEARCOM project.

The relative delay is the difference in processing delay between the left and right hearing aids and should also be kept low. A quick calculation provides the order of magnitude of the tolerable relative delay, given the distance between the ears which is approximately 25 cm. In the case of monaural use of hearing aids the relative delay is equal to the processing delay. For binaural use and localization of sounds, a maximum relative delay of 0.05 ms is tolerable for the Hearing aids/PCS system.

<table>
<thead>
<tr>
<th>Delay Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processing delay</td>
</tr>
<tr>
<td>0  &lt; 15ms</td>
</tr>
<tr>
<td>Relative delay</td>
</tr>
<tr>
<td>0  &lt; 0.05ms</td>
</tr>
</tbody>
</table>

**Table 1 Delay requirements**
4.3.2.2 Signal Quality requirements

Current digital assistive hearing devices provide a high level signal processing. For a one-way link the external audio is sampled (or re-sampled) at least at 16 kHz with a maximum of 32 kHz for high quality applications, and that the subsequent processing is done at this sampling rate. A source of degradation arises when several encoding and decoding stages occur in tandem. The encoding and decoding that is done in the WBAN interface should be transparent with respect to the decoded audio in the PCS generic audio receiver. The consequence of multiple encoding and decoding stages will be tested with the GSM and 3G/UMTS mobile telephony codecs as well as with high-fidelity audio codecs.

For a two-way link, which includes wireless exchange, the influence of channel errors from the wireless transmission is added to the degradations for the one-way link. For an error free channel, the encoding-decoding stages should be transparent or near to transparency. Channel errors can cause distortions that are very annoying to a listener and can interfere with subsequent signal processing. Explicit error control by channel coding and careful choice of the source coding method can reduce the effects of channel errors. The choice of source coding method is crucial in this respect as source codecs can be designed to be inherently robust to channel errors. The channel error rate depends on the transmission method with different rates and different types of errors (burst, bit...). Evaluation of the error rates of the wireless link will provide the guidelines for the amount of channel error protection needed in the communication.

For the requirements the system must provide transparent quality for random errors up to an error rate of $10^{-4}$ and the degradation for higher error rates should be graceful. For burst errors it is required that channel errors do not produce distortions that are perceived as uncomfortable by the user.

<table>
<thead>
<tr>
<th>Signal Quality Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Sampling frequency</td>
</tr>
<tr>
<td>o 32 kHz</td>
</tr>
<tr>
<td>• Robustness to random bit error</td>
</tr>
<tr>
<td>o Transparent for $10^{-4}$ BER</td>
</tr>
</tbody>
</table>

Table 2 Signal quality requirements

4.3.2.3 Physical Requirements

The PCS will be a mobile device that easily can be used when on the move (walking, traveling). For being mobile the following requirements should apply:
Physical Requirement

- Power consumption
  - > 16 hours of operation; > 2 days standby
- Size of the PCS
  - Hand palm size (not bigger than a typical PDA)

Table 3 PCS Physical requirements

4.3.2.4 Mobility requirement

The HEARCOM application is based on the availability of continuous mobile communication of the user with many peripherals, terminals and servers where information or applications are used. We can distinguish different parts of this communication link scenario. The personal communication device is the main part of the system; it is the core of all user communications.

4.3.2.5 Communication requirements

4.3.2.5.a WAN (Wide Area Network link)

In the case of a communication scenario where the user of the PCS is connected to other users and servers via mobile networks, the service, available with a guarantee of high quality and continuity, is provided via GSM, GPRS, EDGE or 3G/UMTS networks with low to medium data transmission speed ranging. The service provided is a phone voice service, messaging and data IP communication, with some limitations due to the relative low speed achieved and also the real cost involved. The big advantage is the availability almost everywhere and the guarantee of high quality and continuity of service while on the move.

Table 4 WAN Communication link scenario

In some cases the user of the PCS is connected to other users and servers via local area wireless networks. These networks can be used to distribute high data bandwidth internet service, even if they often offer no guarantee of service (specific hot spots by companies and institutions).

Via the internet PCS users can have access to local services including public distribution of information and announcements together with
connection to servers that will provide services like speech-to-text conversion. The PCS can be used to adapt to other local network specifically for hard of hearing people, such as the “telecoil”, an analog technology that is presently used to distribute content and announcements in a public environment or to connect to a telephone.

4.3.2.5.b PAN (Personal Area Network link)
In the case of medium distance communication scenarios, the link provides connection to the PCS of the user to the different personal domestic equipments via Bluetooth (TV, radio, etc.).

<table>
<thead>
<tr>
<th>PAN: Personal Area Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Medium range link: Home/personal area: max 5 to 50 meter; Private</td>
</tr>
<tr>
<td>• Bluetooth, WiFi</td>
</tr>
<tr>
<td>• Medium complexity, small/medium size, medium power, data, audio, video</td>
</tr>
</tbody>
</table>

Table 5 PAN Communication link scenario

4.3.2.5.c Personal Communication Link (Body Area Network link)
For short distance communication between the PCS and nearby peripheral equipment, like hearing aids or body worn audio equipment, a wireless body area network (WBAN) can be used, this service is provided via a short range low power radio link.

<table>
<thead>
<tr>
<th>PCL: Personal Communication Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Short range link: Body: max 0.3 to 1.5 meter; Private</td>
</tr>
<tr>
<td>• Low complexity, miniature, low-power, data, audio</td>
</tr>
<tr>
<td>• Coolflux (provisionally selected WBAN);</td>
</tr>
</tbody>
</table>

Table 6 PCL Communication link scenario

4.3.2.6 Processing requirements
The processing requirements include the running of assistive audio signal processing in real-time, as well as running wireless communication services.

From the deliverable D-9-3 the minimum CPU performance requirement of the prototype MHA hardware platform was estimated at 300 Millions Instruction Per Second (MIPS) which relate for large part to floating-point processing. In order to reduce power consumption most of PDA devices (as well as soft phones) use currently DSP processor based on fixed-point architecture. Fixed point provides a quite low computational performance for audio processing. That is the reason why the PHS and PCS cannot be put together in the same device. PHS and PCS are planned to fuse? in a
single PDA-sized device with powerful audio processing capabilities and communication interfaces, but for the moment these devices are considered separately.

Concerning PCS device, the computational performance depends on the communication services that will be included. The required communication services are: 1) connection to the mobile network 2) connection to the wireless public-address network and 3) connection via Internet to a client server speech to text conversion service. On this in potential a global or local positioning service can be added when defined necessary.

In order to achieve these goals, three parts need to be distinguished:

- Management of Input/Output (I/O), when wired and/or wireless connections operate,
- Man Machine Interface (MMI) which has to be handled without disturbing the real-time process,
- Related signal processing tools, including speech/audio codecs, filtering, and other required treatment.

In the case of the I/O management, the PCS device has to use reader/writer process into I/O buffers in order to receive or send data streaming. The requirement of computational performance is then reduced to the minimum. The most important aspect is to handle the interrupts of different tasks efficiently.

The MMI process required more processing powerful which can be maximize by ten MIPS. As previously, the most important aspect is to handle efficiently the graphic interface and the potential interrupts of graphic tasks without to reduce performance of real-time process (i.e. mobile phone communication).

The most critical part for computational performance is obviously the signal processing, which can include many high computational consumption tools such as audio codecs, speech enhancement, speech-to-text tools, etc. For that, it is more convenient to expect more than 300 MIPS according to the selection of these algorithms.

In conclusion, the requirement for the processing aspect of the PCS demonstrator platform is about 400 MIPS. This calculation excludes the PHS floating point processing which for the PCS demonstrator will be implemented on a separate device.
4.3.3 Hardware requirements

The hardware oriented requirements are summarized in the following table. This is in fact based on the specifications of current PDA’s and the announced devices that will be marketed in the next few months.

The following table summarizes the minimal requirements for the development of first version of PCS device. It must have enough processing capabilities and memory and an acceptable autonomy taking into account the current state of technology. It should also incorporate wireless connectivity through GSM/GPRS, WLAN (802.11b) and Bluetooth, as required.

<table>
<thead>
<tr>
<th>Features</th>
<th>Minimum requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor</td>
<td>DSP &gt;200MHz or Xscale &gt;400Mhz</td>
</tr>
<tr>
<td>Memory</td>
<td>128 MB total memory</td>
</tr>
<tr>
<td>Expansion Slots</td>
<td>SD Slot or Mini-SD slot or CF slot</td>
</tr>
<tr>
<td>Power</td>
<td>Battery - Removable/Rechargeable 1800 mAh Lithium-ion.</td>
</tr>
<tr>
<td>Autonomy</td>
<td>200 hours on standby, 4 hours on talk time</td>
</tr>
<tr>
<td>Wireless Technologies</td>
<td>GSM/GPRS, Wi-Fi 802.11b, Bluetooth 1.1, IrDA, UMTS optional</td>
</tr>
<tr>
<td>I/O ports</td>
<td>1 headset connector, IrDA and USB</td>
</tr>
<tr>
<td>Display</td>
<td>TFT 64k, Resolution (W x H): 240 x 320, 2.8”-in diagonal</td>
</tr>
<tr>
<td>Audio</td>
<td>Integrated microphone, receiver, speaker and one 3.5 mm stereo headphone jack.</td>
</tr>
</tbody>
</table>
| Pointing device and keyboard | Touch Screen and Stylus  
|                           | Keyboard: Removable or not 26 alpha-numeric keys |
| Operating system          | Microsoft® Windows® Mobile™ 5.0 or equivalent |
| Dimensions and Weight     | 80 x 25 x 150mm maximum size              |
|                           | 300 g maximum weight                      |

**Table 7 Hardware requirements for PCS platform**

4.3.3.1 Add on Interfaces

The candidate device for PCS platform can have not all the required specifications or features for the targeted applications. Add-on interfaces can be used to fit the device to support add-on hardware. These can be WBAN-dongle; GPS-dongle, Mobile-phone card, etc. The following add-on interfaces are found for PDA’s:

- **USB Connection;** In general PDA’s will support slave USB 1.1 variants; In future also USB 2.0 master variants will become available.

- **SD/SDIO/mSD;** Interfaces for memory extensions but also used to extend hardware (SDIO). Several vendors provide devices.
• CF: Compact flash card. Also suitable for extending hardware functionalities. Several vendors provide devices.

In addition some PDA vendors support private interfaces; which should be avoided in prototyping the PCS, in order to be compatible with the mainstream developments of handheld devices.

An overview on the main characteristics of these interfaces, including some examples, is given in appendix 4.

For the PCS demonstrator preferably the USB and at least the SDIO interface should be supported, allowing flexibility on the design of the WBAN dongle (WP-8 requirement).

### 4.4 Overview on available mobile PDA's

The Personal Communication System will be a mobile device applying following requirements:

- Easily portable for all day wear in pocket or in a small bag: small size and low weight (like a mobile phone or a personal digital assistant);
- Powered for at least 12 hours of operation without battery change (mix of applications);
- Including mobile phone and wireless communication functionalities;
- Enabling high-level signal processing capacity;
- Graphical display for presentation of visual information, and audio interfacing.

The ideal platform is so a PDA or a mobile phone platform that includes PDA functionality (Smart Phone).

### 4.4.1 PDA market

Worldwide shipments of personal digital assistants (PDAs) totaled 3.4 million units in the first quarter of 2005, which is an increase of 25 percent year-on-year [61].

Most important suppliers in order of PDA shipments are:

1. Research in Motion’s (RIM) BlackBerry,
2. PalmOne PDA,
3. Hewlett Packard, the third biggest manufacturer of PDA's saw shipments rise 4.4 percent. The company's market share now stands at 17.6 percent. In addition,

4. Nokia (after re-entry into PDA market).

The world market share by operating system is the following:

![Market share of PDA Operating Systems](image)

**Figure 3 Market share of PDA Operating Systems**

Both classifications are according to the Gartner market study.

It is recognized that Smart-phones, mobile phones with PDA-like abilities, will curtail growth in the sales of PDA's without telephony capability in the near future, as they include more of the same functions.

4.4.2 Overview on available mobile PDA’s

Table 8 provides a list of available mobile PDA devices with some required specifications.

The high range PDA market seems to be moving towards a device capable of handling the maximum wireless technologies with equivalent devices. In addition to Mobile 2nd and 3rd Generation Mobile, WIFI, Bluetooth, and in some cases GPS are supplied in this range of devices as standard communication equipment in addition to the USB and IrDA connectivity.

The processing capabilities of every PDA is difficult to compare with the others. Indeed, if the processor is identical or of the same family, then the performances are comparable, but if the processor is different (DSP, ARM, GPP) then the performances are difficult to compare and the capability depends on the software to be implemented. On the other hand, even if the same type of processor is compared, the processing capability depends also on the internal architecture (parallel or not), and if the useful software is implemented using parallel architecture properties.
For indicative information, most speech coders and speech processing techniques are used to be implemented in DSP processor. Some of them are already available for ARM processor or GPP, but it is still rare. So the implementation of signal processing software is often optimized on DSP, which is more appropriated for mathematic computations, but it isn’t optimized on ARM or GPP, which are more dedicated to the HMI process or telecom software.

### 4.5 Description of selected PDA for demonstration

The PDA iPAQ HP6340 was selected because it meets the minimum requirement for the PCS system. In particular the device meets all required wireless connections and the optimal processing power of 210 MIPS added with 400 MIPS by DSP is expected to support the PCS requirements.

However the device has been discontinued by HP and other more recent PDA’s provide equivalent and higher specifications. The new alternative with equivalent functionalities is the IPAQ H6915 [2][7]. This device provides in particular more processing power, incorporates a camera, memory expansion up to 1GB and GPS functionality.
The PDA market changes very quickly every year, so the final selected device for demonstrator and later PCS versions are expected to exceed the required specifications.

4.5.1 Technical specifications

The Pocket PC HP 6340 is the smallest portable device with three wireless modes (GSM/GPRS, WLAN and Bluetooth) for any voice and data communications. This PDA provides the largest range of high data rate wireless connections for Internet, e-mail, text messaging and radiotelephony:

- Integrated Wi-Fi: offering a high data rate wireless access to Internet, to e-mail and to data sources.
- Bluetooth compatible to facilitate the synchronization of data, printing, and the sharing of files between other personal devices.
- Connection to GPRS networks which supply a permanent access to the data services and to the Internet, all over the world.

The processor OMAP 1510 is very fast and it allows to manage very well GPS applications, receivers GPS Bluetooth, while phoning with a Bluetooth headset.
### Processor, Operating System and Memory

<table>
<thead>
<tr>
<th>Operating System Installed</th>
<th>Microsoft® Windows® Mobile™ 2003 software for Pocket PC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor</td>
<td>Texas Instruments OMAP™ 1510</td>
</tr>
<tr>
<td>processor speed</td>
<td>168 Mhz</td>
</tr>
<tr>
<td>Standard Memory</td>
<td>64 MB SDRAM; 64 MB ROM</td>
</tr>
</tbody>
</table>

### System features

<table>
<thead>
<tr>
<th>External I/O Ports</th>
<th>USB Desktop cradle/charger</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supported Flash Memory Cards</td>
<td>SD Memory Card</td>
</tr>
<tr>
<td>Compatibility</td>
<td>SDIO</td>
</tr>
<tr>
<td>Display</td>
<td>Transreflective type TFT colour with LED Backlight; Number of Colours: 64K colour (65,536 colours) 16-bit; Touch-sensitive screen; Resolution (W x H): 240 x 320</td>
</tr>
<tr>
<td>Display size</td>
<td>3.5-in diagonal</td>
</tr>
<tr>
<td>Display type</td>
<td>TFT</td>
</tr>
<tr>
<td>Keyboard</td>
<td>Removable 26 alpha-numeric keys with 2 application buttons, and send/end buttons – (Included in the box)</td>
</tr>
<tr>
<td>Mouse/Pointing Device</td>
<td>Touch-sensitive display for stylus; 4 programmable application launch buttons</td>
</tr>
<tr>
<td>Wireless capability</td>
<td>Yes</td>
</tr>
<tr>
<td>Wireless Technologies</td>
<td>Integrated GSM/GPRS, Integrated WLAN 802.11b, Integrated Bluetooth® wireless technology; IrDA (SID supported)</td>
</tr>
<tr>
<td>Internal Audio</td>
<td>Integrated microphone, receiver, speaker and one 3.5 mm stereo headphone jack, MP3 stereo (through audio jack and speaker)</td>
</tr>
<tr>
<td>Power supply type</td>
<td>AC Input: 100~240 V, 50/60 Hz, AC Input current: 0.3 A max, Output Voltage: 5Vdc (typical), Output Current: 2A (typical)</td>
</tr>
<tr>
<td>Size (l x p x h)</td>
<td>75x18.7x119mm</td>
</tr>
<tr>
<td>Weight</td>
<td>190 g</td>
</tr>
</tbody>
</table>

### Table 9 HP iPAQ Pocket PC HP6340 specifications

This phone represents latest generation of Pocket PC phones - with built-in Wi-Fi that can be used on top of Bluetooth and GPRS connectivities. This phone is an interesting solution for convergent device.

**Some technical information for the iPAQ:**

1. HP iPAQ phone is based on WANDA platform: WANDA is a reference design made by TI. HP iPAQ phones are practical implementation of WANDA reference design. This design consists of TI OMAP 1510 processor, Bluetooth chip BRF 6100, WiFi chip TNETW 110B and chipset for GSM/GPRS - TCS 2100.

2. HP iPAQ phone is powered by TI OMAP 1510 processor at speed of 168 MHz. This processor consists in reality of 2 processors: ARM-925 processor (to run Pocket PC applications) and TMS 320C55x (a signal processor that takes care of communication jobs).

3. HP iPAQ phone has built-in application for wireless connectivity called "iPAQ Wireless", and this application makes it possible to
switch on and off various wireless connectivity: GSM/GPRS, Wi-Fi and Bluetooth. It is possible to use all three in parallel.

4. The iPAQ phone uses Bluetooth software from WIDCOMM. This offers many Bluetooth profiles and is reliable.

4.5.2 OMAP processor

The selected iPAQ HP6340 applies the OMAP5910 processor. This is a dual-core architecture that combines the command and control capabilities of the TI-enhanced ARM™ 925 processor with the high-performance and low power capabilities of the TMS320C55x™ DSP core. This processor is designed to maximize system performance and minimize power consumption.

The OMAP processor is schematized below:

![Figure 5 OMAP processor](image)

The Processor consists of:

- One 150-MHz, 210 MIPS TI-enhanced ARM9 processor core for control and high-level OS including 128-kByte Internal Dual-Port SRAM
  - 24 KB Cache
  - Data and Instruction MMUs
  - 32-bit and 16-bit instruction sets
• One 150-MHz, 400 MIPS C55x™ DSP core for accelerating applications with:
  o 24 KB Cache
  o 160 KB SRAM
  o 32KB ROM
  o Hardware Accelerators for video algorithms
• a DSP/BIOS™ bridge between the ARM and the DSP
• Peripherals and on-chip Resources:
  o 192 KB shared SRAM
  o Two 16-bit memory interfaces for SDRAM and Flash
  o Nine-channel system DMA controller
  o LCD Controller
  o USB 1.1. Host and Client
  o Eight serial ports plus three UARTs
  o Eight Timers
  o Real Time Clock
  o Keyboard interface
  o 18 GPIO pins

4.5.3 Development kit for OMAP processor

For the IPAQ HP6340 the Innovator development platform is used. This platform consists of:

• Breakout board with Ethernet port, mouse and keyboard
• User interface module:
  o 240*320 LCD touch screen – 16 bit, 64K color
  o 4-wire touch screen controller
  o Dual RS232 port interfaces
o Microphone, stereo speakers and stereo out

o USB host and client port

- Processor module:
  o OMAP 5910 processor
  o 32M SDRAM, 32M Flash
  o JTAG accessible

---

**Figure 6 Innovator kit platform**

---

4.5.3.1 Other Development tools

The target platform is linked to a PC for running the software development tools. These PC-based development tools include:

- For the DSP C55x side, the development tool is Code Composer 2.2.

- In the ARM side, the Operating system used is the WINCE OS 4.1 (Microsoft) and the tool used to develop is Visual embedded.

- Active Sync is used to download in the ARM the code developed on Visual Embedded tool.

- To debug DSP application on the platform, a JTAG USB link is used.
4.6 Communication services

4.6.1 Wide Area Network (WAN) communication

For communication scenario the user of the PCS is connected to other users and servers via mobile networks, the service, available with a guarantee of high quality and continuity, is provided via GSM, GPRS, EDGE or 3G/UMTS networks with low to medium data transmission speed ranging. The service provided is a phone service, messaging and data IP communication in mobile application.

4.6.1.1 Mobile phone function

The first application is the access to the normal mobile telephone system for the hearing impaired people to receive and generate phone calls whatever the conditions are.

The idea is to introduce an intermediate device between callers using a remote powerful computer able to convert in real time the sound. The second option, and the best one, is to use directly a PDA integrating the mobile phone function. Most PDA have the phone functionality.

In the current PDA’s the GSM or 3G function is either integrated in the processor (DSP or GPP) or it is present in a specific chipset. The goal is to get signal input, before the GSM or 3G function in order to apply specific useful processing. The main idea is to integrate audio and speech enhancement processing to improve and personalize the communication.

4.6.1.2 Wireless data application

Recent mobile devices allow data wireless application (file transfer, client/server application,..) with increasing quality of service. If GPRS (and also more recent improvements) represents a current useful solution, the new 3G technology provides better functionalities.

4.6.1.3 Wireless LAN application

Most of PDA terminals provide WIFI and/or Bluetooth technologies in their functionalities. These solutions allow wireless communication with local network applications.

- EDGE (Enhanced Data rates for GSM Evolution) provides data transfer rates significantly faster than GPRS,
- 3G wireless networks providing data transfer rates of up to 384Kbps. future 3G networks are expected to reach speeds of more than 2Mbps,
- GPS Localization service optionally available on PDA’s.
4.6.2       WBAN

Further information about the Wireless Body Area Network or WBAN is available in the deliverable D-8-4. However, some technical points can be repeated.

WBAN is usually linked to sensor applications where a high number of devices placed near the body can provide status information to a central system.

In HearCom WBAN will be employed to convey multi-channel audio information to devices like microphones or headsets with sufficient performance as to configure a high quality hearing aid system. In this sense it is a high throughput system that needs to be transmitting during long periods of time. This makes the HEARCOM WBAN especially challenging. However, it is thought that if this can be achieved then not only will there be a new range of technologies for hard of hearing users, but that the wireless technology produced will also be useable by a wider market, for example wireless headphones for personal entertainment systems.

4.6.3       Personal Area Network (PAN)

Personal Area Network describes the range of communications that could go from some meters for Bluetooth to a maximum of 100 m (for WIFI).

This type of network permits the communication of the PCS to the external IP world by a combination of fixed LAN that reaches to the WIFI or BT access points. The last leg is then covered by the mentioned PAN technologies.

The type of application taking advantage of PAN can range from the general IP World access (Web surfing, VoIP, Internet Services ) to access to the Home appliances and entertainment world (TV, HiFi).

The relatively low cost access to the Broadband Internet via WIFI is to be differentiated from the Mobile network access. In the PAN case the access is confined to the access range of the PAN network with very limited roaming capabilities.

4.7       Hardware extensions

Additional hardware functionalities can be added to the iPAQ using existing interfaces. Most of PDA provides Input/Output slots such as compact flash or SD slots, USB which allow external additional plug-ins.

For the PCS demonstrator all required communication hardware will be part of the standard iPAQ hardware. Optionally the following add-on hardware will be possible:
• WBAN dongle: by interface on SDIO or USB. This dongle will be part of WP8. For testing of the WBAN including the new codecs for hearing devices the PHS will be used. When possible the WBAN will also be ported to the PCS for demonstration. For this the WBAN dongle can be applied.

• GPS dongle for localization. Optionally when required for accurate localization for public announcement services. (not planned to be included).
5 Assistive communication services

The PCS system may support several applications and services. For the scope of the project the following assistive services will be defined:

1. Public Announcement and Address Systems (WPA),
2. Speech recognition system (client-server ASR),
3. Personal Hearing System (PHS).

In addition the PCS will have the following supportive services:

1. Advanced assistive information display for use with ASR and WPA,
2. Phone service adapted to ASR.

5.1 Public Announcement and Address Systems

Public announcement services are used in public transport stations, as train, underground, bus stations, and airports. Events with high participation can use as well this service as Public Address Systems: churches, schools, concerts, sport, conferences workshops and so on, in order to broadcast the event content, issue warnings announcements and provide additional or useful information.

Following the line of increasing possible services offered through public announcement, it is getting more and more popular to use this in malls, big shops, cinemas, theater... etc. in order to provide additional information (besides information on screens, ads, and so on).

There is another way of providing public announcement service: by means of the broadcast and dissemination possibilities that mobile and itinerant networks present.

Once the Public Announcement System makes use of the available technology like wireless communication systems and Internet to distribute the information a number of additional services becomes possible including more efficient catering for user disabilities.

The various types of public announcement services belong to the two types of architectures and scenarios described in deliverable D9.3: Hearing in the communication society (section 4, [8]). As the nature of messages on both types of announcements is different, the system which has to process them has different architecture.
In Category 1 (direct microphone audio input), the incoming audio stream is received with microphones integrated in a hearing aid. The audio stream reaches the PHS directly via WBAN from the hearing aid. Therefore, the acoustic feedback from the receiver to the microphone is a very critical issue and requires special headsets (hearing aids fitted to the individual ear) and advanced signal processing. In the Category 2 (Communication link audio input-via PCS) the incoming audio stream is not received via the local hearing aid but transmitted via a communication channel (PCL), provided by the PCS to the PHS. The audio stream is received by the PCS via PCL. This situation does not depend critically on very low delay, and acoustic feedback does not take place.

In figure 7 it is shown the common architecture suitable for the examples commented below.

![Figure 7 Architecture for Public Announcement System](image)

**Figure 7 Architecture for Public Announcement System**

In order to explain the roles involved in a public announcement situation, it is going to be described as an example:
--- a person arriving at the airport---.

When a person arrives at a public transport station, he can receive information of different types: audio, text and vibration (alarm information).

Audio information comes from loudspeakers (i.e. flight announcements), audible alarms, audible adds).

Text information can be of different types:

- Dissemination information, like advertisement about offers (flight ticket booking, hotel reservations, holiday, mobile operator ads, and so on). This type of information is received by means of GPRS, WAP, Bluetooth and WiFi, on mobile devices and PDAs with this characteristic.

- GPS information: once the person has arrived at the station, it can be activated the GPS functionality, and it can be received information about the person location, and the nearest place of interest (tourist places, bus stations, metro stops, taxi stops, even some events that are going to take place, and the place in which they are going to be celebrated). Of course, the user will have the option of configuring some special filters, in order to avoid the information’s he does not want to receive.

- Other localization techniques particularly useful for inside Building and underground are the Wireless mobile local and personal networks configured by Cell Phone, WIFI and BT.

Vibration alarms could be also configured, depending on user profile in order to accommodate to the user characteristics and preferences. Commonly, alarms are associated to text or voice messages. The announcing of the alarm warning or new message or phone call could be a vibrating device.

Vibration alarms are also useful in other kind of environments, as can be: home, workplaces, etc. The alarm can be associated to a little text which includes the core of the message showed in the screen.

Public announcement services can be configured as Ambient Intelligence systems that provide messages in different formats to users depending on the user needs, profile, and global and local positioning. The user access to the information should be as focused as possible on the needs without manipulation or profile modification made by the user. For example if the user has booked for an specific flight the rest of flight announcements will not have access to him.
The location of the user is determined by the different systems available from Satellite based geo-localization to cell based (Mobile network) or to local radio beacon.

The user profile can be defined initially based on explicit user questionnaires and evolve based on ongoing experience and user decisions.

These factors, user profile and location will be used to filter and format the actual information that arrives to the user in the various forms.

The basic backbone of the service will be the digital telecommunication networks Mobile, WIFI, Bluetooth and others for wireless communication and Ethernet for fixed communication.

Due to the variety of legacy systems now available for the hearing handicapped person in the Public Announcement Services (Loop, IR, FM, etc...) there is a need to adapt those to the new digital technologies providing an interface and entry gate to the digital network as can be seen in Figure 8.

**PUBLIC ANNOUNCEMENT SYSTEM**

<table>
<thead>
<tr>
<th>Possible Involved Technologies</th>
<th>Services</th>
<th>Usable Outputs</th>
<th>Output Medium</th>
</tr>
</thead>
<tbody>
<tr>
<td>WiFi</td>
<td>Spoken voice (microphone)</td>
<td>Amplified Sound</td>
<td>Hearing Aid</td>
</tr>
<tr>
<td>Bluetooth (Beacon)</td>
<td>*Airports PA, metro/subtrain stations PA</td>
<td>*Malls, workshops, conferences, events, schools PA</td>
<td>Speakers</td>
</tr>
<tr>
<td>WBAN (user ↔ PCS)</td>
<td>Synthetic voice (microphone)</td>
<td>Written Text</td>
<td>Headphones</td>
</tr>
<tr>
<td>GPRS</td>
<td>*The same as previous</td>
<td>*Possibility of sending dissemination messages associated to the recorded messages</td>
<td>Screen</td>
</tr>
<tr>
<td>GPS</td>
<td>Electrical Equipment</td>
<td>*To include TV/AlertRadio</td>
<td>*Ability to manipulate font, size, type and colour, *Rolling/substitext</td>
</tr>
<tr>
<td>FM IrDA Loop</td>
<td>Broadcast Dissemination Messages (text/audio/image)</td>
<td>Vibration</td>
<td>*Ability to filter and configure reception of dissemination messages</td>
</tr>
</tbody>
</table>
5.1.1 Public Announcement demonstrator

A PA system demonstrator is proposed with the most important following elements as shown in Figure 9:

User:

- Equipped with PCS and hearing aids connected WIFI/GPRS. With specific profile of capabilities and interests.

Airport Gate:

- Identifies users as they approach via WIFI beacons. Sends information specifically adapted to users interests in the adequate format including voice text and images.

Figure 9 Demo Proposal PA system
5.2 Speech recognition system (client-server)

5.2.1 Phone Service

One of the important services for the hearing-impaired person is the access to the normal telephone system. To be able to receive and generate phone calls and to be supported when due to hearing impairment difficulties the conversation between two partners becomes impossible.

The main idea of the speech recognition service is to introduce an intermediate device between callers using a remote powerful computer able to convert in real time the sound only information of a typical telephone call into multimodal information (sound plus text and/or phonemes).

Once the call is established if one of the call participants demands the service of HEARCOM ASR their voice information will be available to the remote computer together with the personal profile of the participants (caller identification, technical characteristics of the call, language and accents, presentation requirements etc.). The central HEARCOM computer will by means of automatically generated voice and text, introduce the service to the participants and controls the conversion of individual voice utterances to text. This will involve feedback control to both speakers requiring if necessary the verification that the sentence has been correctly converted to text and the repetition request to the speaker of not well understood sentences.

The resulting text is sent to the hearing-impaired user via the available data channels for its presentation in a variety of possible formats that include the full translated text in a user profile dependent and configurable presentation.

This service can be the initial step and can later be extended to other application like phone services for the elderly, on the fly translation for foreign speakers etc..

5.2.2 Functionality

The functionality of speech recognition as a client-server application enables the communication between two parties, of which one party is hearing impaired and can be assisted by speech to text conversion. The hearing impaired person uses a PDA device to connect to the server that provides this service (the IP of the server is specified in the text box seen below):
The person using the PDA is able to receive acoustic and text data from the person with whom he/she is having a session. The other person speaks, their utterance is sent to the server, where automatic speech recognition is performed, and then the audio signal together with the recognized text (and optionally of the phoneme output) are sent to the PDA of the hearing impaired person.
The incoming text is appended on top of the existing text, so that the most recent recognized utterance is on top. Furthermore, the last utterance can also be re-heard by pressing the Play button. This will be a useful option for a hearing impaired person. The Record button is used so that the user of the PDA can as well send speech signal to the other party.

The functionalities of speech recognition system are described in the table below.

<table>
<thead>
<tr>
<th>Functionality for ASR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario 1: Dictation</td>
</tr>
<tr>
<td>Scenario 2: Spontaneous speech</td>
</tr>
<tr>
<td>Languages : Greek, English</td>
</tr>
<tr>
<td>Output of recognizer: words, confidence scores, time stamps</td>
</tr>
<tr>
<td>Output of recognizer: phonemes</td>
</tr>
<tr>
<td>Call characteristics: environmental noise suppression</td>
</tr>
<tr>
<td>Vocabulary selection: general, medical, legal</td>
</tr>
<tr>
<td>Speech-text ASR guidance (ASR instructions, feedback)</td>
</tr>
<tr>
<td>Speech-text operator mode</td>
</tr>
</tbody>
</table>

**Table 10 Functionalities of ASR System**

5.2.3 ASR Phone Service

The ASR phone application configuration consists of a number of elements as shown in the following figure.

- User equipped with a PCS terminal
- Correspondent user, calling from a mobile or a fixed line telephone via analog (a/b) or digital ISDN lines as well as VoIP lines
- Server with associated Telephone interface card.
5.2.4 Client-Server Functions

The server performs different functions:

- User management
- Phone Interface
- ASR/TTS Functions
- ASR Feedback
- Server-Client

As observed in the following figure:
5.2.4.1 User Management.

HEARCOM users are identified by means of a user registration and user/password mechanism. All communication is established via a secured interface. The individual user profile is not stored at the server but at the personal user terminal PCS.

A set of predefined profile characteristics is sent at the beginning of the call to initiate all server functions with the corresponding profile parameters.

5.2.4.2 Phone Interface.

The phone interface module is in charge of configuring the Audio Card in the server to connect to the different networks available for transmission of the voice information and signaling.

5.2.4.3 Client Server Logic.

The client-server handles the communication between client (PCS) and server and maintains the flow of data information.
5.2.4.4 Feedback Module.

This module controls the dialog between the Server and the different call participants by means of rolling text, and automatic voice messages.

The objective is to introduce the service to the participants, handle the initial parameterization of the call and assure the quality of the ASR service by giving instructions and presenting the proposed ASR result for confirmation. This could be performed providing automatic answers on the phone by the server via Text to Speech converter with sentences like “you just said... if is it OK press 1”.

5.2.4.5 Automatic Speech Processing (ASR) and (TTS).

Includes the two main speech processing activities: Automatic Speech Recognition (ASR) and Conversion from Text to Speech (TTS).

ASR Automatic Speech Recognition transforms Coded Audio Voice as received via the voice channels into text and/or phonemes and using different ARS techniques.

Based on the defined caller profile that specifies Language, Accent, Vocabulary, Conversion speed, Codec, Phoneme/Text etc. It produces a stream of text corresponding to the best approximation of the received audio stream together with the percent of estimated correctness. The text stream may also contain special emotion and tempo markings.

The Text to Speech module converts text to voice. This will be used to communicate with the callers via voice, in order to give automatically instructions and feedback on the ASR operation for the phone conversation.
5.3 Personal Hearing System

The Personal Hearing System (PHS) is targeted to be an integrated service of the Personal Communication System (PCS). PCS and PHS together are planned to become a mobile-phone-sized (PDA-sized), with powerful audio processing capabilities and communication interfaces. It has to be noted that this is a long term target that does require an intermediate step. The previously released deliverable D9-3 describes a short term standalone PHS based prototype/demonstrator as well as gives an outline on future integration possibilities. In the following we want to give a short outline. For further details the reader is kindly asked to refer to the original document.

The PHS enables powerful, exchangeable signal processing services on the PCS without limiting these to a defined set of processing algorithms. The PHS bases on a software framework named Master Hearing Aid (MHA) that defines an API for easy algorithm development and integration. The algorithm themselves are the focus of e.g. WP5. Executing the results of WP5 on the PHS will be a straight forward and easy task. The PHS accepts audio-streams from the PCS that may originate from a variety of sources. These sources could be on the PCS itself (e.g. integrated MP3-Player or...
Voice-Mail), from the Body Area Network (BAN) (e.g. Hearing Aid with Personal Communication Link), or from an external source (e.g. mobile phone or public announcement system). Regarding the different audio sources we differentiate two scenarios.

1. Situations where the incoming audio stream is received with microphones integrated in a hearing aid

2. Situations where the incoming audio stream is not received via the local hearing aid microphones (e.g. telephone communication)

![Figure 15 Two scenarios of PHS-usage](image)

The left side of the figure is depicting a first scenario audio-stream processing. In this scenario the PHS-service functions as a hearing aid extension. Depending on the latency of the link the PHS can offer control parameter processing for the hearing aid only or audio signal enhancement additionally. With a short latency link (delay < 10-15 ms) the PHS is able to execute signal processing without the user to be able to notice a visual discrepancy. The signal stays synchronous to the moving lips of a speaker. With a larger delay the PHS is still useful to perform complex feature extraction algorithms on the audio stream. These algorithms can be used to calculate parameter sets for the hearing aids’ signal processing. An example would be scene recognition where the PHS identifies a situation with “speaker in the front” and sets the hearing aids’ microphones to a directional receiving pattern while it opens up the angle during a “concert/theatre” audio situation.

The right side of the previous figure is depicting a second scenario audio-stream processing. This scenario does not make use of hearing aid microphones. The delay of the PCL is not an issue. It is even possible to replace the hearing aids by a regular (Bluetooth) headset. Together with
the PHS the low budget headset becomes a hearing aid replacement that is expected to be very attractive for hearing impaired that refuse to obtain a regular hearing aid but need assistance during e.g. phone calls. An external audio-signal from a cell-phone (could be part of the PCS) or public announcement system is being processed by the PHS to enhance signal-quality and comprehensibility.

A prototype for each of the two scenarios will be implemented. The scenario one prototype includes a USB-Soundcard to simulate a fully featured hearing aid. At this point no hearing aid is available that offers a powerful link from the aid to the PHS-prototype. For this reason it was chosen to use dummy headsets (including three microphones and one speaker without signal-processing) instead and simulate the hearing aid signal-processing together with the PHS-services on the same prototype HW-platform. The missing WBAN wireless link with the hearing aids will also be simulated for evaluation purposes.

This figure gives an overview on the scenario one prototype. The top of the three rows shows the final integrated target system as it should be. A PDA-like device includes PCS- and PHS-services and connects via a PCL to
hearing aids. The middle row shows the software architecture of the complete system: PCS-services, PHS-services, wireless link and hearing aids with integrated signal processing. The hearing aid internal signal-processing can be seen as “smaller” PHS-versions. The bottom row shows the targeted prototype hardware. Only the PCS-services stay on a PDA-like device, while the PHS-services are implemented on a notebook together with a wireless link simulator and a hearing aid simulator. The later requires the aforementioned USB-audio device for analogue to digital (and vice versa) high quality signal conversion.

This figure gives an overview on the scenario 2 prototype. The target system now includes simple headsets that are widely available on the market. This makes it much easier to implement a prototype that is close to the target system. Since we do not need high bandwidth to include six hearing aid microphones but one set only it is possible to use a wireless, low bandwidth bluetooth connection. Latency is of no concern since the scenario targets phone communication or public announcement systems where visual synchronicity is not required. A very interesting hybrid solution between a bluetooth headset and a hearing aid is the ELI DirX.
Module. By clipping this module to hearing aids with separate audio input it is possible to use hearing aids in the same way as headsets. Unfortunately the ELI device is not good for the scenario one prototype since it has its own microphone and does not enable access to the hearing aids’ microphones.

The two prototypes implement the complete services that are part of the PHS. So we show a complete picture regarding the services of the PHS in this section of the document. The roadmap towards the final integrated version of PCS with PHS in one device is part of deliverable D9-3.

### 5.4 Assistive information display

The PCS terminal display has the following principal display functions:

1. **Terminal control**
   This concerns the human textual/graphical interface to control the functions of the terminal and its applications.

2. **Information**
   This concerns all streams of textual and graphical information provided by the PCS services, like e.g. public announcement message.

The PCS system basically will apply the standard PDA display modes for control and the normal standard PDA applications.

The advanced assistive information display will be a visual information display for application of the bimodal comprehension of speech and text information. This display will present the assistive visual information corresponding to and derived from auditory information like speech.

Visually displaying assistive information that corresponds to the auditory information can improve communication performance in adverse listening situations in which auditive information is imperfect due to for example background noise interference or due to hearing impairment. Presenting assistive visual information corresponding to the speech signal is expected to increase communication performance (The basic research for this will be reported in deliverable D-9-6). The bimodal application will be researched and prototyped for two PCS services.

- Announcements in which partly intelligible auditory information is supplemented by textual information that is obtained from an Automatic Speech Recognition service (ASR) or directly from the original information source (when textual information is transmitted from the information source).

- Phone conversation in which speech from the connected party is converted to text using a client server ASR system.
In addition and support to the above real-time bimodal mode an off-line recording mode will be made available in the PCS. Recording will include both the available synchronous auditory and textual information. This allows off-line retrieving of information that has not been captured, understood or memorized. The recording mode may be applied to several services such as the announcement service, phone conversations, ASR applications and other to be defined.

5.4.1 Visual display

The visual display will consist of the presentation of text in a continuous or rolling way.

The "rolling" adjective confers the meaning of dynamic service: the test rolls in the screen, discarding old information and updating and presenting the most recent information, while the audio conversation is going on.

The information shown in the display can be plain text or text with special markings that codify the translation success rate or the emotional content of the auditory message. It can also include various types of phoneme representation.

In order to read the visually displayed information, the advance button (cursor), or the scroll will be used. The cursor will be controlled by a hardware scroll device. In the following figure this is shown in indicated circles:

![Figure 18 Rolling Text Elements by Scroll](image)

In order to facilitate the readability of the presented text, the interface of the display has to be clear and display characteristics like the font format, font size, presentation mode, background, and language have to be
optimized and/or adaptable to the preferences of the user. An example of the interfaces is shown in the next figures:

![Image of interfaces showing display options]

**Figure 19 Rolling text Human Machine Interface (example)**

The Human Machine Interface presentation will have the following options:

- Displaying the recognized words,
- Displaying the recognized phonemes,
- Displaying the recognized words and phonemes simultaneously.

The user profile options will include:

- Setting of language,
- Format options (for each mode: text, phoneme):
  - font type and font size,
  - color, contrast,
  - other (e.g. alignment, spacing, etc).
- Scroll mode: Direction of text: vertical or horizontal mode.
5.4.2 Phoneme presentation

For the presentation of phonemes the following options can be provided:

- **IPA**: International Phonetic Alphabet; This representation will require a specific character set that must be made available at the PCS. See Appendix 1.

- **SAMPA**: Speech Assessment Methods Phonetic Alphabet; a computer readable phonetic alphabet that uses 7-bit ASCII characters and is based on IPA. SAMPA is available for many languages. See Appendix 2 for example. See: http://www.phon.ucl.ac.uk/home/sampa. This representation will use (one of) the standard character sets and has no further special requirements for the PCS.

Phoneme reading is not used normally and some training will be required. The device can support this by a specific training mode. In this mode phonemes and words will be displayed simultaneously while hearing the corresponding voice sound.

Reading phoneme-like text is effortful, partly because of the fact it is quite difficult to detect word-boundaries. Listening to distorted and incomplete speech (due to the hearing impairment) and simultaneously reading phonemes and/or words might be hard.

The next figure shows the possibilities of presenting phonemes.
It has to be noticed that in the figure placed on the right side the sentences registered are delineated according to ASR perceived pauses and speaker recognition between the sentences. This will enable the listener to quickly find the corresponding phoneme transcription while listening to the speech.

5.4.3 Caller Information

Caller information can be very useful to the listener, in order to predispose himself to “his caller”. As it has been previously commented, each person has different “talking” characteristics, and the listener could become accustomed to some selected callers. When the caller is calling, a protocol will be initialized in order to get additional information like the name and nationality of the caller and the environmental characteristics (codec type, noisy, indoor, outdoor...).

Furthermore, the topic or subject of the conversation could be presented to the listener at the outset of the call (as it is used in emails, like subject Key words). This option could be initiated by the caller at the beginning of the conversation. Conversation topics could for example be: Commercial, Appointment, Personal conversation, Work conversation, Sport conversation. This possibility is even more interesting if the ASR word-database is indexed according to these topics. These possibilities have to be further discussed and developed conveniently and will require further study taking into account the ethical review conclusions.
Caller information can be hidden activating the Caller Info button placed on the right side.

5.4.4 User profiles

The user can establish its own preferences (personal preferred parameters) in the corresponding profile file in the PDA.

The main idea is to have a flexible service, related to configuration parameters, in order to fulfill all personal requirements.

The default parameters will be provided for the initial trials according to the preferences from technologic and human interface partners and consolidated after the corresponding user trials.

5.4.5 Functionalities advanced information display

In this section the possible functionalities of the rolling text display service are described. As it has been introduced in a user profile section, there are many possibilities and it is important to discuss all of them in order to define the functionalities that will be implemented.

For this objective we also need the recommendations of people working in hard-of-hearing clinics, hearing experts, speech and acoustic experts, and technical partners.
### Functionality

<table>
<thead>
<tr>
<th>Functionality</th>
<th>Nice to have</th>
<th>Must</th>
<th>Demo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode 1: Presentation of Speech-Text transcription</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mode 2: Presentation of Speech-phoneme transcription</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mode 3: Presentation of both modes</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mode 3.1: expanded presentation of both modes (mode 1 in more than middle)</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mode 3.1.1: numbering of sentences presented</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mode 3.2: intercalary modes (normal-phoneme, normal-phoneme…)</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horizontal presentation</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vertical presentation</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Save text conversation into a file</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extract parameters of the conversation (phone numbers, dates, timetables, addresses, Proper Names…)</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Text format changing: choose the characteristic fonts preferred (size, color, type…) for both modalities</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Record and Reproduce conversations</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Navigation through the text by means of cursor or scroll buttons</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Caller information:</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Name</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Call characteristics:</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Environment: chosen among options of a list (noisy, indoor, outdoor)</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Key words: to be told by caller</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Topic/Subject: to be told by caller</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nature: to be chosen among options of a list (Familiar, Work, Personal, Medical, Appointment)</td>
<td>?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Table 11 HMI Rolling text display functionalities

## 5.5 Service Interactions

PCS is at the center of a number of devices: Hearing aids on one side that communicates via PCL and PCS to the rest of available communication services like audio phone systems (fixed, mobile or IP based), signaling (alarms of diverse origin: Public Announcement system, Home systems), entertainment (Home Video or Audio systems) Public Announcement or Address Systems conveying audio information.

Based on this a number of functions can be implemented in PCS that will involve some other devices: PCL, Hearing Aid or any Remote Sound Server. These functionalities can be automatic or could require the user interaction via PCS keyboard or any other means.

### 5.5.1 Interaction with hearing aids

- Hearing aid status report:

  Reports on status of main parameters (battery, malfunctions, diagnostics) and current profile and activity.
• Hearing aid load profile:
  Will load new hearing aid profile
• Hearing aid mode change
  o Hearing aid activate: Full Hearing aid function
  o Local mode: Direct feed from hearing aid microphones
  o Hearing aid Remote mode: Only communication from external communications services, phone, audio, ASR.
  o Hearing aid silence: Complete silence.
  o Hearing aid standby: Saves battery only accepting alarms and event signals for hearing aids
  o Hearing aid test mode: Initiates HA self test, no user feedback.
  o Hearing aid calibrate: Initiates calibration test of hearing aid with user feedback.

5.5.2 Interaction with PCL
• PCL status report: Reports on status of main PCL parameters (binary error rate, packet loss, delay, interferences pico net activity and elements, battery, malfunctions, diagnostics) and current profile and activity. Services available.
• PCL load profile: Will load new PCL profile.
• PCL service discovery and subscription.
• PCL mode change.
  o PCL activate: Full PCL function
  o PCL silence: Complete silence.
  o PCL standby: Saves battery only accepting alarms and event signals
  o PCL test mode: Initiates PCL self test, no user feedback.
  o PCL calibrate: Initiates professionally calibration test of PCL with user feedback.
5.5.3 Interaction with mobile phone

This will correspond to the usual Mobile Phone functionalities:

- Directory
- Messages,
- Personalization
- Tools
- Events
- WAP
- Applications.

In the case of the PCS the Mobile phone is integrated in the equipment.

5.5.4 Interaction with VoIP phone

The interaction with VoIP phone will supported when a Internet Access Network via WLAN or directly via USB to and Ethernet connected PC cable is available. Thus it corresponds to the usual VoIP Phone functionalities as achieved by a private access to VoIP services or commercial (Skype, Yahoo).
6 Software Components

In Table 12 is shown a list of the Software components proposed to operate around the HEARCOM PCS.

Where convenient Open Source Software has been selected. However for PDA’s at the moment the MS Windows Mobile is dominant in the market in spite of inroads from the Linux world.

<table>
<thead>
<tr>
<th>Elements</th>
<th>OS</th>
<th>Software</th>
<th>Components</th>
<th>Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCS</td>
<td>Windows Mobile</td>
<td>C#, visualStudio.net</td>
<td>PCL WiFi</td>
<td>VoIP, Data Internet Global and local positioning</td>
</tr>
<tr>
<td></td>
<td>C# Visual studio.net</td>
<td>PDA</td>
<td>PCL GPRS/GSM/UMTS</td>
<td>Phone, Data, Internet, VolP</td>
</tr>
<tr>
<td></td>
<td>C# Visual studio.net</td>
<td>GPS</td>
<td>Global and local positioning</td>
<td></td>
</tr>
<tr>
<td>PHS</td>
<td>Linux</td>
<td>C</td>
<td>ASR</td>
<td>Core DSP</td>
</tr>
<tr>
<td>HEARCOM Server</td>
<td>Linux/Windows Server</td>
<td>Java/J2EE</td>
<td>TomCat/JBoss</td>
<td>Web Service &amp; Server Applications</td>
</tr>
<tr>
<td></td>
<td>Linux/Windows Server</td>
<td>Java/J2EE</td>
<td>LDAP or Database</td>
<td>User Management</td>
</tr>
<tr>
<td>HEARCOM Public Announcement Server</td>
<td>Linux/Windows Server</td>
<td>Java/J2EE</td>
<td>Apache</td>
<td>Application Server</td>
</tr>
<tr>
<td></td>
<td>Linux/Windows Server</td>
<td>Java/J2EE</td>
<td>TomCat/JBoss</td>
<td>Web Service &amp; Server Applications</td>
</tr>
<tr>
<td></td>
<td>Linux/Windows Server</td>
<td>Java/J2EE</td>
<td>LDAP or Database</td>
<td>User Management</td>
</tr>
</tbody>
</table>

Table 12 Table of Software Components

6.1 Extension and adaptations for Software

By taking into account the overall system, the included peripheral Software (WBAN), which is planned to be implemented in ULP technology, could be coded in low level C or Assembler language. The PCS will be implemented in a PDA, the PHS which will be implemented in a Tablet or Portable PC and the servers giving service to the different applications planned.

The PCS platform needs to integrate new modules to be able to deal with these applications. Therefore, multiple tasks have to be conducted simultaneously. For instance, the PCS device must be able to handle in the same time wireless communications with PHS or hearing headset, via PCL, and with WPA server or ARS HEARCOM server, via WPAN or WLAN. It has
to integrate also text display software to provide useful information to the end-user. Such capabilities are not integrated in current commercial devices.
7 Dissemination and exploitation

In a first step, concerning the dissemination aspect, the development of the PCS platform and its use in assistive communications services can be largely presented to the professional, technical community and general public by means of:

- HEARCOM Web page,
- Hearing Impaired and Audiology forums and Congresses,
- Professional and Scientific Press (Journals),
- Technical Fairs (e.g. CeBit in cooperation with PDA manufacturers).

Then, with the Wireless public address demonstration or the ASR demonstration, and especially if they lead to positive results, this prototype will be of interest for firstly PDA manufacturers, and also public place organizations (airport, cinema, congress center, etc.). The ASR application is moreover of great interest for telephone service providers.

Potential relations with other projects can be obviously established. Lot of technical, national or European projects deal with wireless communication and its potential applications and services. For instance, in the field of digital cinema, the EDCINE [10] project studies the wireless transmission of audio and video streaming in real theatres. This work could be an opportunity to define suitable protocol or common means to realize interactions with the HEARCOM PCS platform. This aspect can also be a first opportunity for the standardization aspect, by defining new protocols.

The definition of WPA and ASR scenarios lead to the utilization of wireless communication standard. A goal of the HEARCOM project is not only to use existing wireless standard but also to involve in the modification or improvement of these standards or also to define, if required, new suitable standards. Further information could be given in the next step of our work.

8 Ethical Issues

The storage of user profiles and recording of conversational information may have ethical issues that should be analyzed in a next phase. For the demonstration system the ethical issues will not be a limiting factor.
9 Conclusions

The Personal Communication System (PCS) is a device onto which a collection of services and technologies designed to improve communication can be hosted. Key to the success of these proposed services will be the connections between the PCS and other technologies such as mobile phones, public address systems and the client-server speech to text conversion system.

This deliverable covers essentially the selection and the definition of the PCS platform and the description of the communication services. These services will focus on the interaction and connection of the PCS to mobile phones, public-address and the client server speech text conversion system.
10 References

[1] Gartner market study on PDA’s
[3] FP6–004171 HEARCOM, D-8-1: Requirements specification of personal link and related application services
[8] FP6–004171 HEARCOM, D-9-3: Hearing in the communication society
[9] FP6–004171 HEARCOM, D-8-3: Analysis of existing and innovative assistive applications
Appendix 1: International Phonetic Alphabet

THE INTERNATIONAL PHONETIC ALPHABET (revised to 2005) © 2005 IPA

<table>
<thead>
<tr>
<th>CONSONANTS (PULMONIC)</th>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Dental</th>
<th>Alveolar</th>
<th>Postalveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Uvular</th>
<th>Pharyngeal</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plosive</td>
<td>p</td>
<td>b</td>
<td>t</td>
<td>d</td>
<td>t</td>
<td>q</td>
<td>k</td>
<td>g</td>
<td>q</td>
<td>g</td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trill</td>
<td>B</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tap or Flap</td>
<td>V</td>
<td>f</td>
<td>f</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td>Ø β v</td>
<td>θ ð s z</td>
<td>s z</td>
<td>c j</td>
<td>x y ch h</td>
<td>h f i</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral fricative</td>
<td>l h</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td>u l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral approximant</td>
<td>l l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Where symbols appear in pairs, the one to the right represents a voiced consonant. Shaded areas denote articulations judged impossible.

<table>
<thead>
<tr>
<th>CONSONANTS (NON-PULMONIC)</th>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Dental</th>
<th>Alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Uvular</th>
<th>Pharyngeal</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clicks</td>
<td>Ø</td>
<td>d</td>
<td>t</td>
<td>d</td>
<td>t</td>
<td>q</td>
<td>k</td>
<td>g</td>
<td>g</td>
</tr>
<tr>
<td>Voiced implosive</td>
<td>b</td>
<td>d</td>
<td>t</td>
<td>d</td>
<td>k</td>
<td>g</td>
<td>q</td>
<td>g</td>
<td>?</td>
</tr>
<tr>
<td>Ejercitives</td>
<td>p'</td>
<td>t'</td>
<td>d'</td>
<td>k'</td>
<td>g'</td>
<td>q'</td>
<td>g'</td>
<td>i'</td>
<td>i'</td>
</tr>
</tbody>
</table>

OTHER SYMBOLS

<table>
<thead>
<tr>
<th>W</th>
<th>Voiced labial-vowel aspirant</th>
<th>Z</th>
<th>Alveolar-velar diphthongs</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>Voiced labial-vowel approximant</td>
<td>J</td>
<td>Voiced alveolar lateral flap</td>
</tr>
<tr>
<td>F</td>
<td>Voiced labial-palatal approximant</td>
<td>H</td>
<td>Voiced palatal approximant</td>
</tr>
<tr>
<td>S</td>
<td>Voiceless palatal approximant</td>
<td>G</td>
<td>Voiceless glottal approximant</td>
</tr>
</tbody>
</table>

DIACRITICS

Diacritics may be placed above a symbol with a descender, e.g. ʃ.

VOWELS

<table>
<thead>
<tr>
<th>Front</th>
<th>Central</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>i y</td>
<td>i u</td>
<td>uu u</td>
</tr>
<tr>
<td>e ø</td>
<td>ø ø ø</td>
<td>ø o</td>
</tr>
<tr>
<td>Open</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e ø</td>
<td>e ø ø</td>
<td>e ø</td>
</tr>
<tr>
<td>Close</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e ø</td>
<td>ø o ø</td>
<td>o ø</td>
</tr>
</tbody>
</table>

SUPRASEGMENTALS

<table>
<thead>
<tr>
<th>Primary stress</th>
<th>Secondary stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>jœfluæk</td>
<td>jœfluæk</td>
</tr>
<tr>
<td>,,</td>
<td>,,</td>
</tr>
<tr>
<td>Long</td>
<td>Half-long</td>
</tr>
<tr>
<td>Extra-short</td>
<td>Minor (foot) group</td>
</tr>
<tr>
<td>Major (intonation) group</td>
<td>Syllable break</td>
</tr>
<tr>
<td>Linking (absence of a break)</td>
<td></td>
</tr>
</tbody>
</table>

TONES AND WORD ACCENTS

<table>
<thead>
<tr>
<th>Level</th>
<th>Contour</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>Rising</td>
</tr>
<tr>
<td>low</td>
<td>Falling</td>
</tr>
<tr>
<td>rising</td>
<td>High</td>
</tr>
<tr>
<td>falling</td>
<td>Low</td>
</tr>
<tr>
<td>down</td>
<td>Rising down</td>
</tr>
<tr>
<td>down</td>
<td>Falling</td>
</tr>
<tr>
<td>up</td>
<td>Global rise</td>
</tr>
<tr>
<td>up</td>
<td>Global fall</td>
</tr>
</tbody>
</table>

Definition of PCS platform and assistive communication services
Appendix 2: SAMPA computer readable phonetic alphabet

Consonants

The standard English consonant system is traditionally considered to comprise 17 obstruents (6 plosives, 2 affricates and 9 fricatives) and 7 sonorants (3 nasals, 2 liquids and 2 semivowel glides).

With the exception of the fricative /h/, the obstruents are usually classified in pairs as "voiceless and "voiced", although the presence or absence of periodicity in the signal resulting from laryngeal vibration is not a reliable feature distinguishing the two classes. They are better considered "fortis" (strong) and "lenis" (weak), with duration of constriction and intensity of the noise component signalling the distinction.

The six plosives are p, b, t, d, k, g:

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Word</th>
<th>Transcription</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>pin</td>
<td>pIn</td>
</tr>
<tr>
<td>b</td>
<td>bin</td>
<td>bIn</td>
</tr>
<tr>
<td>t</td>
<td>tin</td>
<td>tIn</td>
</tr>
<tr>
<td>d</td>
<td>din</td>
<td>dIn</td>
</tr>
<tr>
<td>k</td>
<td>kin</td>
<td>kIn</td>
</tr>
<tr>
<td>g</td>
<td>give</td>
<td>gIv</td>
</tr>
</tbody>
</table>

The "lenis" stops are most reliably voiced intervocally; aspiration duration following the release in the fortis stops varies considerably with context, being practically absent following /s/, and varying with degree of stress syllable-initially.

The two phonemic affricates are tS and dZ:

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Word</th>
<th>Transcription</th>
</tr>
</thead>
<tbody>
<tr>
<td>tS</td>
<td>chin</td>
<td>tSIn</td>
</tr>
<tr>
<td>dZ</td>
<td>gin</td>
<td>dZIn</td>
</tr>
</tbody>
</table>
As with the lenis stop consonants, /dZ/ is most reliably voiced between vowels.

There are nine fricatives, \( f \ v \ T \ D \ s \ z \ S \ Z \ h \):

\[
\begin{array}{lll}
\text{f} & \text{fin} & \text{fIn} \\
\text{v} & \text{vim} & \text{vIm} \\
\text{T} & \text{thin} & \text{TIn} \\
\text{D} & \text{this} & \text{DI}s \\
\text{s} & \text{sin} & \text{sIn} \\
\text{z} & \text{zing} & \text{zIN} \\
\text{S} & \text{shin} & \text{SI}n \\
\text{Z} & \text{measure} & \text{"meZ@} \\
\text{h} & \text{hit} & \text{hIt}
\end{array}
\]

Intervocally the lenis fricatives are usually fully voiced, and they are often weakened to approximants (fricationless continuants) in unstressed position.

The sonorants are three nasals \( m \ n \ N \), two liquids \( r \ l \), and two sonorant glides \( w \ j \):

\[
\begin{array}{lll}
\text{m} & \text{mock} & \text{mQk} \\
\text{n} & \text{knock} & \text{nQk} \\
\text{N} & \text{thing} & \text{TIN} \\
\text{r} & \text{wrong} & \text{rQN} \\
\text{l} & \text{long} & \text{lQN} \\
\text{w} & \text{wasp} & \text{wQsp} \\
\text{j} & \text{yacht} & \text{jQt}
\end{array}
\]

**Vowels**

The English vowels fall into two classes, traditionally known as "short" and "long" but, owing to the contextual effect on duration of following "fortis" and "lenis" consonants (traditional "long" vowels preceding fortis consonants can be shorter than "short" vowels preceding lenis consonants), they are better described as "checked" (not occurring in a stressed syllable without a following consonant) and "free".

The checked vowels are \( I \ e \{ Q \ V \ U \):

\[
\begin{array}{lll}
\text{I} & \text{pit} & \text{pIt} \\
\text{e} & \text{pet} & \text{pet} \\
\{ & \text{pat} & \text{p{t} \\
\text{Q} & \text{pot} & \text{pQt} \\
\text{V} & \text{cut} & \text{kVt} \\
\text{U} & \text{put} & \text{pUt}
\end{array}
\]
There is a short central vowel, normally unstressed:

\[ \@ \text{another} \quad @"\text{nVD}@ \]

The free vowels comprise monophthongs and diphthongs, although no hard and fast line can be drawn between these categories. They can be placed in three groups according to their final quality: i: eI aI OI, u: @U aU, 3: A: O: I@ e@ U@. They are exemplified as follows:

- i:  ease  i:z
- eI  raise  reIz
- aI  rise  raIz
- OI  noise  nOIz
- u:  lose  lu:z
- @U  nose  n@Uz
- aU  rouse  raUz
- 3:  furs  f3:z
- A:  stars  stA:z
- O:  cause  kO:z
- I@  fears  fI@z
- e@  stairs  ste@z
- U@  cures  kjU@z

The vowels /i:/ and /u:/ in unstressed syllables vary in their pronunciation between a close [i]/[u] and a more open [I]/[U]. Therefore it is suggested that /i/ and /u/ be used as indeterminacy symbols.

\[ \begin{align*}
\text{i} & \quad \text{happy} & \quad "\text{h\{pi} \\
\text{u} & \quad \text{into} & \quad "\text{Intu}
\end{align*} \]

NOTES.

1. Notational variants. Differently from the notation set out above:

   (i) It is possible to transcribe English long vowels without using length marks, thus /i u 3 A O/. This is phonemically unambiguous, although it does remove the option of restricting the symbols [i u] to the use just described, for the phonemically indeterminate weak vowels.

   (ii) The symbol /E/ is quite widely used in place of /e/ for the vowel of "pet".

   (iii) In an older notation, now no longer in general use, paired short and long vowels were transcribed using the same vowel symbol with
and without length marks, thus /i/ in "pit", /iː/ in "ease"; /o/ in "pot", /oː/ in "cause".

2. Additional symbols. For some purposes and some varieties of English it is useful to give explicit symbolization to the glottal stop and/or the voiceless velar fricative:

```
?     network             neʔw3ːk
x     loch               1Qx
```
Appendix 3: External networks and Internet

GPRS:

GPRS (General Packet Radio Service) is a packet based wireless communication service that offers data rates from 9.05 up to 171.2 Kbps and continuous connection to the Internet for mobile phone and computer users. GPRS is based on GSM communications and complements existing services such as circuit switched cellular phone connections and the Short Message Service (SMS). GPRS represents the bridge between 2G and 3G mobile telecommunications and is commonly referred to as 2.5G. GPRS implementation requires modification of existing GSM networks, because GSM is a circuit switched technology while GPRS is packet oriented. GPRS enables packet data (the same as is used by an Ethernet LAN, WAN or the Internet) to be sent to and from a mobile station - e.g. mobile phone, PDA or Laptop. WAP and SMS can also be sent using GPRS and individuals working with GPRS need to learn and understand how the mobile stations, the air interface, network architecture, protocol structures and signaling procedures must be modified.

GPRS offers much higher data rates than GSM and can be combined with 3G technologies such as EDGE (Enhanced Data-Rates for GSM Evolution) to give even higher bit-rates. It offers many benefits for customers and network operators: such as volume (rather then time) dependent billing and more efficient use of network resources.

By taking into account to the worldwide delay in implementing 3G solutions such as CDMA and UMTS the demand for GPRS is still growing.

When the GPRS is connected to the IP network, it behaves as a IP network. The mobile GPRS terminal has an IP address. This address is either public attributed during the subscription, or dynamic supplied during the connection. The GPRS network in itself is formed in fact by the SGSN-GGSN:

- The node of SGSN service (Serving GPRS support Node) is connected with one or several BSS (Base Station Subsystem: BSC+BTS), and the router manages terminals in a given zone (see following figure);

- The node of bridge GGSN (Gateway GPRS support Node) is connected with one or several data networks, it is the router which allows input packets, coming from external data networks, to be forwarded to the SGSN of the receiver and to send output packets towards the adequate data network.

For the external IP network, the GGSN is seen as a router.
EDGE (Enhanced Data rates for GSM Evolution) provides data transfer rates significantly faster than GPRS or HSCSD. EDGE increases the speed of each timeslot to 48 kbps and allows the use of up to 8 timeslots, giving a maximum data transfer rate of 384 kbps. In places where an EDGE network is not available, GPRS will automatically be used instead. EDGE offers the best that can be achieved with a 2.5G network, and will certainly be replaced by 3G.

3G Technology

3G wireless networks are capable of transferring data at speeds of up to 384Kbps. Average speeds for 3G networks will range between 64Kbps and 384Kbps, quite a jump when compared to common wireless data speeds that are often slower than a 14.4Kb modem. 3G is considered high-speed or broadband mobile Internet access, and in the future 3G networks are expected to reach speeds of more than 2Mbps.
Appendix 4: Add-on Interfaces for PDA

USB Connection

USB (Universal Serial Bus) is the most flexible solution to connect new hardware like digital joysticks, scanners, digital speakers, digital cameras or a PC telephone to a computer. USB makes adding peripheral devices extremely easy. With USB-compliant PCs and peripherals, you just plug them in and turn them on. USB makes the whole process automatic. It's like adding instant new capabilities to your PC. You never need to open your PC, and you don't need to worry about add-in cards. And you don’t need to shut down and restart your PC to attach or remove a peripheral. Just plug in the new device and the PC automatically detects the peripheral and configures the necessary software.

PDA's uses USB connections, and the first application is connection, for synchronization or to exchange data, with the PC. They include the capability to recharge the PDA power of via USB cable.

USB 1.1 uses two speeds: LowSpeed (1.5 Mbps) and FullSpeed (12 Mbps) not to be confused with USB 2.0.

USB 2.0 is the next generation USB interface technology. This technology is also referred to as Hi-Speed USB.

You can transfer data between the computer and peripherals 40 times faster than original USB. Hi-Speed USB technology offers transfer rates up to 480Mbps compared to USB 1.1 devices, which transfer at speeds of 12Mbps.

USB 2.0 will become the standard for future peripherals using the USB interface connection. However, it will not replace original USB technology in all products. Next generation USB allows users to transfer data faster between peripherals and the computer. Users are also able to connect up to 127 devices to a single computer by adding hubs to create more USB 2.0 ports. It is still rare to have USB 2.0 available in a current PDA.

SD/SDIO/mSD

A SD Card (Secure Digital Card) is an ultra small flash memory card designed to provide high-capacity memory in a small size. SD cards are used in many small portable devices such as digital video camcorders, digital cameras, handheld computers, audio players and mobile phones. In use since 1999, SD Memory Cards are now available in capacities between 16 Megabytes and 8 Gigabyte. An SD card typically measures 32 x 24 x 2.1 mm and weighs approximately 2 g.

SDIO (Input/Output) card is an interface that extends the functionality of devices with SD card slots. A variety of SDIO cards are being developed,
such as a Camera, Bluetooth, GPS, and 802.11b/g. Of course, the host device for the SD I/O Card will also be able to use the SD Memory Card.

After the success of the SD Card, the miniSD Memory Card was developed to meet the demands of the mobile phone market. The miniSD Card provides the same benefits as the SD Card, but is smaller than the original SD Card. MiniSD Cards are typically found in many newer mobile phones and PDA’s with features such as built-in digital cameras, downloading and games; basically the mobile phones where the miniSD can meet the requirements for increased data storage. MiniSD cards are 21.5 x 20 x 1.4 mm and generally provide 16MB to 256MB of storage, but goes up to 2 GB nowadays.

The plan for the HEARCOM project is to implement the WBAN as an add-on SDIO card.

SDIO (Secure Data Input/Output) card is then an interface that extends the functionality of devices with SD card slots. A variety of SD I/O cards are being developed, such as the SD Camera card, SD Bluetooth card, SD GPS card, and SD Mobile Communication card.

![Figure 24 SDIO cards with external functions](image)

CF card

In the same way the Compact Flashcard is an interface that extends the functionality of devices with CF card slots. A variety of CF cards are being developed, such as the CF Camera card, CF Bluetooth card, CF GPS card, and CF Mobile Communication card.
Appendix 5: Public Announcement Demonstrator Activities

The Public Announcement System proposed will focus on a system capable of handling the voice and data communications of a Transport Public Announcement System:

This scenario covers Transport Systems and in particular the Gate Information system in an Airport.

The Hearcom user is able to receive in the PCS wireless environment: voice and data announcements corresponding to the user profile; this includes Warnings, Gate related announcements and other airport announcements. These announcements can be filtered by specific User Profile including: travel plan, language, interest areas etc.
Below follows an overview on activities to develop a functional Demonstrator for the PA System. This overview is for reference only. The actual Demonstrator will be reduced to contain the essential PCS modules and concentrate on user aspects in particular on retrieving and presentation of information. The reduction will be detailed in the next period.

10.1.1 Overview of activities

Studies:

Contemplates the performance of studies on the relevant issues for the HEARCOM Demonstrator this activity will use the result of previous or simultaneous deliverables D8.1,D8.2,D8.3,D8.4 D9.1,D9.2 D9.3, D9.7,D9.8.

These studies collect the relevant information for the demonstrator implementation and perform if required additional detailed studies. The amount of manpower reserved takes into account the maturity of the findings of the previous HEARCOM studies and the foreseen difficulties.

Analysis:

During Analysis the Detailed Specification Definitions of the different elements of the PA System including viability analysis are generated. Detailed specs should include:

- Architecture
- Requirements
- Protocols and interfaces.
- Data model.
- Test plan

Design:

During Design the Specification is implemented in a concrete Module Block diagram

Should Include:

- GUI and HMI
- Data base creation files
- Module definition
• User documentation

**Implementation:**
Reflects the actual coding of the Modules.

**Local test:**
Elements are tested individually simulating the rest of the system.

**Integration and Test.**
The complete system is put together in a lab environment and tests are performed verifying the complete functionality and throughput. Network performance is simulated in a lab environment.

**Final System Test:**
The complete system is installed and tested in a real environment.

**Demonstrator Trial:**
Contemplates the support effort for each demonstrator trial.

The total required manpower for a full demonstrator is 40 to 60 person-months. This is not feasible as for the actual demonstrator only about 10 person-months will be available. For reducing the effort it is proposed that part of the modules (i.e. at server side) will be emulated by simulation. This will be detailed in next phase.