ETSI EG 202 191 V1.1.1 (2003-08)

ETSI Guide

Human Factors (HF);
Multimodal interaction, communication and navigation guidelines

Reference

DEG/HF-00029

Keywords

ICT, disability, multimode, Design for All,
interaction, interface, diversity, user

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Contents

Intellectual Property Rights 5
Foreword 5
Introduction 5
1 Scope 6
2 References 6
3 Definitions and abbreviations 7
3.1 Definitions 7
3.2 Abbreviations 8
4 Lessons from stakeholder consultation 8
5 Current initiatives for multimodal accessibility 9
6 Usage scenarios 10
7 Multimodality for reduction of social exclusion 10
7.1 The social exclusion of disabled people 10
7.2 Monomodality of information presentation and social exclusion 11
7.3 Multimodality as compensation for impairments 11
7.4 Alleviating social exclusion 12
7.5 Benefits of multimodality for everybody 12
8 Multimodality for improving accessibility 12
8.1 What is Accessibility? 12
8.2 Assistive technologies 13
8.3 Physical disabilities 13
8.4 Low vision 14
8.5 Blindness 15
8.6 About hearing disabilities 15
8.7 Communication and language disabilities 16
9 Human factors guidelines and recommendations for multimodal interaction 17
9.1 Metaprinciples for the design of multimodal systems 17
9.2 Design principles for multimodal systems 18
9.3 Implementation principles 21
10 Conclusions and recommendations for future work 21
Annex A (informative): Stakeholder consultations 23
A.1 Questionnaire(s) 23
A.2 Records of consultations 30
A.2.1 User Consultation, Perth, 20th June 2002 30
Annex B (informative): Applying the design guidelines to a real world scenario 38
Annex C (informative): Current initiatives (technology roadmap) 46
C.1 World Wide Web Consortium (W3C), Web Accessibility Initiative (WAI) 46
C.1.1 Judy Brewer - Director, Web Accessibility Initiative, International program Office 46
C.2 Media Access Group at WCBH, National Centre for Accessible Media (NCAM) 46
C.2.1 Geoff Freed, Project Manager 46
C.2.2 Gerry Field, Technology Manager 46
C.3 Critical Computing Research Group, MIT Media Lab 47
C.4 IBM Research Labs 47

C.4.1 Shari Trewin....47
C.4.2 Stéphane Maes - Manager, Mobile Speech Solutions, human language technologies....47
C.4.3 Jay Murdock - Industry Solutions Lab, & Teresita Abay-Krueger - DISU Event Manager, Industry Solutions Lab....47
C.4.4 Jennifer Lai - Speech Interface Design, Pervasive Computing....48
C.4.5 Rick Kjeldsen - Vision-based user interfaces....48
C.4.6 Vicki Hansen....48
C.4.7 Medical Automation Research Centre....48

C.5 Georgia Tech....49

C.6 TRACE Centre, University of Madison....49

C.7 Microsoft....50

Annex D (informative): Bibliography....51

History....53

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (http://webapp.etsi.org/IPR/home.asp).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This ETSI Guide (EG) has been produced by ETSI Technical Committee Human Factors (HF).

Introduction

The eEurope vision is of an information society that releases human and economic potential to improve productivity and the quality of life in Europe. Technology developments are expected to open up significant economic and social opportunities. From new technologies are expected to spring new services, applications and content that will create new markets and provide the means to increase productivity and hence growth and employment throughout the economy.

By 2005 Europe is expected to have modern online public services providing e-government (see note), e-learning and e-health to a population of citizens working in a dynamic e-business environment. As an enabler for these services there will be widespread availability of broadband access at competitive prices supporting a secure information infrastructure.

NOTE: eEurope 2005: An information society for all, An Action Plan to be presented in view of the Sevilla European Council, 21/22 June 2002 http://europa.eu.int/information_society/eeurope/news_library/documents/eeurope2005/eeurope2005_en.pdf.

Improved access for disabled people is an objective of the eEurope initiative. In October 2001 the Council of the European Parliament adopted a resolution on e-inclusion and a further resolution in March 2002 stating that Member States should speed up their efforts in implementing the "Web Accessibility Initiative" guidelines. The vision of eEurope sees ICT as an opportunity for disabled people, and other vulnerable groups, to gain equality of access to participation in society not as just another means by which they are to be disenfranchised or excluded.

There are clearly social, regulatory and commercial drivers pushing large corporate ICT users to demand accessibility in the systems they purchase. It will also be increasingly necessary for them to retrospectively add accessibility to existing technologies in use for day-to-day commercial activities. Providers of ICT services are facing increasing demands for accessible services to meet the needs of older consumers, and also from legislative pressure from both national and EU actions.

There is an opportunity here for accessibility technology to add value to new wave multimodal information technology. A technology capability in this area could be used for product differentiation, technology provision and support could be offered to application vendors, or accessibility services offered out through public networks to third parties for resale within their own commercial service offerings. ICT providers need clear guidance on the human issues for the design of this technology from which they can build commercial product that is inclusive of the broad diversity of people who need to make use of it in Europe.

1 Scope

The present document identifies key issues, solutions and actions for multimodal interaction, communication and navigation at the user interface with ICT systems and terminals. It specifically addresses the usage context of transactional interactions for independent living e.g. on-line expenses forms and diaries (employment); on-line shopping and banking (home); on-line gaming and travel agency (leisure); learning delivery systems (education). To add to previous guidance provided by HF such as EG 202 116 [2] and TR 102 068 [3].

The present document focuses on the dynamics of multimodal transactions/user dialogues for the full diversity of users of ICT systems and terminals. It identifies how simplifications, translations, sensory transpositions, or other presentation or content manipulations of a multimodal transaction can be used to improve ease of access for people with sensory, motor or cognitive impairments to telecommunications products and services. It reports the results of consultation with users and user groups to identify the areas of transactional interactions that are currently providing a barrier to ICT access. To overcome these barriers a structured set of design and implementation guidelines is presented.

The present document emphasizes opportunities for simple generic human factors solutions that are commercially attractive to network operators and equipment providers for delivery as a sustainable revenue generating activity. These are open information and communications technologies to consumers who might otherwise be excluded. In addition it makes recommendations for the specific actions that need to be taken in this area to overcome barriers and expand use. Where possible this is supported by illustrations or examples.

The recommendations, guidelines and principles given here have been produced from a "design for all" perspective. While they have been produced out of observation and consultation with disabled people they have been constructed to be of value in making multimodal dialogues easier and more effective for all users not just those involved in their construction. The present document does not address the special needs of young children for access to ICT, which is dealt with in TR 102 133 [4].

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication and/or edition number or version number) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies.

Referenced documents which are not found to be publicly available in the expected location might be found at http://docbox.etsi.org/Reference.

[1] CEN/CENELEC Guide 6 (2002): "Guidelines for standards developers to address the needs of older persons and persons with disabilities".

[2] ETSI EG 202 116: "Human Factors (HF); Guidelines for ICT products and services; "Design for All"".

[3] ETSI TR 102 068: "Human Factors (HF); Requirements for assistive technology devices in ICT".

[4] ETSI TR 102 133: "Human Factors (HF); Access to ICT by young people: Issues and guidelines".

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply:

accessible design: design focussed on principles of extending standard design to people with some type of performance limitation to maximize the number of potential customers who can readily use a product or service

NOTE 1: See [1], [2] and [3].

NOTE 2: Accessible design is a subset of universal design. Terms such as design for all, barrier-free design, inclusive design, transgenerational design are used similarly but in different contexts.

NOTE 3: Design for all is more commonly used in Europe. It refers to designing mainstream products and services to be accessible by as broad a range of users as possible. It can be achieved through one of three ways:

a) by designing products, services and environments that are readily usable by most users without any modification;

b) by making them adaptable to different users (adapting user interfaces); and

c) by having standardized interfaces to be compatible with special products for people with disabilities.

NOTE 4: Barrier-free design is more commonly used in codes and standards documents, and often in reference to the removal of barriers in buildings, whether physical or sensory.

assistive technology device: device used by a disabled person to prevent, compensate, relieve or neutralize any resultant handicap and which has the ability to interface to an ICT device

Design for All (DfA): design of products to be usable by all people, to the greatest extent possible, without the need for specialized adaptation

ICT device: device for processing information and/or supporting communication which has an interface to communicate with a user

impairment: problem in body function or structure such as a significant deviation or loss

NOTE 1: See [1], [2] and [3].

NOTE 2: Impairment can be temporary or permanent, slight or severe and can fluctuate over time.

NOTE 3: Body function can be a physiological or psychological function of a body system; body structure refers to an anatomic part of the body such as organs, limbs and their components.

multimodal: adjective that indicates that at least one of the directions of a two-way communication uses two sensory modalities (vision, touch, hearing, olfaction, speech, gestures, etc.)

multimodality: property of a user interface in which:

a) more than one sensory modality is **available** for the channel (e.g. output can be visual or auditory); or

b) within a channel, a particular piece of information is represented in more than one sensory modality (e.g. the command to open a file can be spoken **or** typed).

navigation: at a high level, navigation information is concerned with the following elements: where am I (and how did I get here), where can I go to, and how do I get there?

EXAMPLE: The navigation bar besides word processors indicates the location where you are in the document, and whether you can go up or down, or both. Moving the block up or down will take you in that direction. Especially when navigating in information spaces, it is not always clear where you are (e.g. in a tree structure), how you can get where you want, etc.

pleasurenomics: the study of human pleasure in the design of products and services and the methods for improving user experience as result of an improved understanding of this aspect of human interaction with products and services

universal design: design of products and environments to be usable by all people, to the greatest extent possible, without the need for adaptation or specialized design

NOTE: See [1], [2] and [3].

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

ADL	Activities of Daily Living
AT	Assistive Technology
ATM	Automatic Teller Machine
CSS	Cascading Style Sheets
DfA	Design for All
DVD	Digital Versatile Disc
EU	European Union
GRP	Generation Research Program
GUI	Graphical User Interface
HTML	HyperText Markup Language
HWZ	Human Studies Centre
ICT	Information and Communication Technologies
LMU	Ludwig-Maximilian University
MARC	Medical Automation Research Centre
MIT	Massachusetts Institute of Technology
NCAM	National Centre for Accessible Media
PC	Personal Computer
PDA	Personal Digital Assistant
SALT	Speech Application Language Tags
SMIL	Synchronous Multimedia Interaction Language
STF	Specialist Task Force
STQ	ETSI technical committee for Speech, Transmission planning, and Quality of service
SUI	Speech User Interfaces
SVG	Scalable Vector Graphics
TDD	Telecommunications Device for the Deaf
TV	TeleVision
HF	ETSI technical committee Human Factors
UAG	Universal Accessibility Gateway
URL	Universal Resource Locator
WAI	Web Accessibility Initiative
WAP	Wireless Application Protocol
W3C	World Wide Web Consortium
Xforms	A GUI toolkit for X-Windows
X-Windows	A hardware-independent network-based window system developed at MIT
XML	eXtended Markup Language

4 Lessons from stakeholder consultation

The majority examples of solutions to accessibility issues focused on the requirements of blind, visually impaired or deaf users. The requirements of cognitively or language impaired users were particularly poorly represented. This reflects the fact that there is insufficient knowledge of the requirements and the potential solutions for these users.

The accessibility of transactional services is a combination of the terminal and the content. It did not appear to simply be a question of access to the input or display from a PC terminal but dependent on an interaction between terminal capability and the design of the content being accessed.

The key trend highlighted by the stakeholders was the shift of focus from the presentation of embedded content to the holding of content in a delivery independent form. This trend is being promoted as a result of the convergence of communications, entertainment and information processing technologies. A number of stakeholders are preparing to deliver seamless interaction with transactional sessions irrespective of location of the user or the device that is being used. An example would be a phone call with a sales person via a telephone or smart phone could spawn a transaction that could be suspended and continued on a TV interface or could be conducted simultaneously on a PDA and a public kiosk. As this evolution progresses, the information relevant to the transaction, the service content, will need to be held in a form that allows it to be presented in an appropriate manner on a range of devices, perhaps presented in a variety of media, and interacted with in a variety of interaction modalities.

This approach addresses the fundamental interactional barrier currently encountered by users with disabilities, that of inaccessible media optimized for delivery on a specific platform using specified modalities. In order to successfully complete transaction presented in this way, the content needs to be transduced into an alternative media and the device adapted to ensure that usable interaction modalities are available. This solution is rarely completely successful. As a result of the changing climate of legislation promoting accessibility for all, service providers are investigating potential accessibility strategies. The delivery independent holding of content is emerging as a potentially attractive solution.

The records of the consultation sessions with users can be found in clause A.2.

5 Current initiatives for multimodal accessibility

The principle current trend in transactional interaction addresses the seamless handover or maintenance of a transaction across different platforms simultaneously or over time. The goal is to be able to engage in interactions on a variety of platforms, and then to continue the same interaction on other platforms as appropriate. For example, an enquiry about train times and availability could be commenced on a mobile phone, continue through a TV set based web browser, and then completed at a station kiosk as a simultaneous interaction involving the kiosk and the mobile phone. It is implicit in this work that the interaction is multimodal, and that the interaction can be conducted in the modalities most suited to the user.

The work brings together a variety of previously disparate technology, including the web and its delivery through visual browsers, voice technologies within telecommunications infrastructures, broadcast technologies such as TV and satellite, and dedicated transaction services such as airline booking or banking systems. For this reason, a variety of existing underlying delivery and interaction technologies are being harmonized and news mechanisms are being developed.

The W3C organization is playing a lead role in the work through its Multimodal Interaction Activity. It is addressing the evolution of the HTML based web delivery with extensions based on Cascaded Style Sheets (CSS), Synchronous Multimedia Interaction Language (SMIL) and Scalable Vector Graphics (SVG). VoiceXML is being explored as a candidate technology for voice interaction. The interaction between the voice and visual interfaces is being explored through the work of organizations such as IBM in their lead of the definition of the Multi-Modal Browser Architecture. Future form based interaction is being proposed to be built upon the Xforms technologies. Fundamental to this whole approach is the drive towards device independent holding of content that should ensure the possibility of seamless mobile transactional interaction.

In parallel to this work, the W3C is also considering the accessibility at the presentation and interaction level in terms of the usability of the client browsers through the work of the User Agent Accessibility Guidelines Working Group.

Whilst the W3C provides a focus for this work worldwide, there are various industrial, standardization and research activities continuing independent, but invariably in cooperation with the W3C. These include:

- Speech Application Language Tags (SALT) work being undertaken by Microsoft.
- Multi-Modal Access activity within Siemens AG.
- The speech interaction activity for 3rd generation mobile within the 3GPP forum.
- Work within ETSI on Distributed Speech Recognition within the Aurora STQ.

The initiatives and activities outlined above have been seeking to address the generic accessibility issues, some organizations have sought to address accessibility in specific contexts. Examples of this are the collaboration between WGBH and MIT address the accessibility of education software and learning systems and the work of the IMS Global Learning Consortium.

Further details of the ongoing work in this domain resulting from study visits by the project team can be found in annex C.

6 Usage scenarios

For the purposes of this study, three primary usage scenarios have been selected as fundamental to independent functioning within society. These are:

***Work*:** Various transactions are now conducted almost exclusively online. Employees are expected to interact with these services in order to participate in the life of the employing organization. These services include the booking of flights and hotels, the submission of expenses claims, and the organization of meetings via electronic diaries.

***Consumer Services*:** The trend to provide "care in the community" as a means of promoting "independent living" has resulted in people moving out of, or not moving into, residential care facilities into domestic or community based dwellings. In many cases, whilst the people are able to function in these dwellings, mobility outside to shops and other services may be difficult. The growth of eCommerce service delivery is potentially an attractive solution, but only if the users are able to conduct transactions in a manner that is accessible to them.

***Leisure*:** As a result of the convergence of broadcast entertainment and computing technologies, interactive multi-party entertainment is emerging as an important class of transactional services. These range from simple WAP based quizzes to more demanding interactive games.

In addition to recognizing the variety of services that are utilized by users, it is important to also recognize that the device through which the transactions are conducted can have an influence on the usability of the transaction. The ranges of devices currently used to interact with transactional services are:

1) Mobile Phones, used via keys.

2) Hand held devices held in the palm of the hand and used via a touch sensitive screen and a pen stylus.

3) Small clamshell systems such as laptop computers or hand held computers. These are used via a keyboard and/or a touch sensitive display.

4) Static system such as a desktop computer or a TV with attached keyboard.

A range of these services and devices were explored with potential users as part of a user consultation exercise in this project. This is reported in clause A.2.

7 Multimodality for reduction of social exclusion

7.1 The social exclusion of disabled people

Disabled people are a group at high risk of social exclusion because of the physical, legal, financial, and attitudinal barriers from society that they face in their everyday life. Social exclusion is a multidimensional phenomenon that is linked not only to employment, income and expenditure, but also to activity status, education, housing, health, and social and interpersonal relations.

Being disabled, so it seems, still goes hand in hand with poverty caused by unemployment. Reliable statistics on unemployment among disabled people are hard to collect due to the fact that European societies employ different definitions of disabilities and use different methods for calculating the unemployment rate. Some evidence, however, suggests that employment rates among disabled people are low, at around 40 % (see note 1) (compared to 65 % of the non-disabled population), and half of those disabled people who have employment earn incomes below half the general population mean (often taken as an indicator of poverty) (see note 2).

NOTE 1: The employment situation of people with disabilities in the European Union (2001). European Commission DG Employment and Social Affairs, Brussels.

NOTE 2: Enduring economic exclusion: Disabled people, income and work (2002) Joseph Rowntree Foundation. http://www.jrf.org.uk/knowledge/findings/socialpolicy/pdf/060.pdf.

Asked about the main causes for social exclusion among people with disabilities, 94 % of the participants of a study conducted in 2001 in several European countries listed the "lack or limited access to social environment and unemployment". 85 % of the respondents (from disability organizations, decision makers, and other stakeholders) listed the "lack or limited access to services", 81 % named the "lack of adequate training" (see note 3).

NOTE 3: Disability and Social Exclusion in the European Union: Time for Change, Tools for Change. Study funded by the European Commission DG Employment and Social Affairs (Project Number VP/2000/008-qurey).

7.2 Monomodality of information presentation and social exclusion

The missing access to environments, services and adequate training seems to contribute to a larger degree to the social exclusion of disabled people than their living in institutions (only listed by 60 % of the sample in the study mentioned above).

The access to the opportunities offered by society is obviously limited if these cannot be reached by persons with impairments or restricted mobility (e.g. persons in a wheelchair or blind persons). A more subtle way of exclusion results from the sensory modalities in which they are presented.

The presentation of information in one modality only excludes people with impairments in that particular sensory modality from taking in that information (a deaf person e.g. has no use for a traditional radio as it provides information only via the auditory sensory mode). This applies to all areas of information presentation including entertainment, education, business transactions as well as conversational services.

Not being able to use a device or service because its input and output channels support one modality only is a serious restriction of one's everyday life. It can mean e.g. that online banking becomes impossible and transactions have to be made personally (which may present other additional barriers).

Information provided in trains or buses by voice or visual indication only seriously restrict those who cannot see or hear.

Similar restrictions exist for large sections of educational programmes that cannot be used by disabled people thus limiting their options for professional promotion. Many jobs are designed around a technology that is not designed in a way as to provide multimodal interfaces - a large portion of working population is working with technology that require seeing and hearing.

The impact of monomodality on a person's social life is also dramatic. A hearing impaired person watching a movie is excluded from the sound, a visually impaired person from the picture (often friends or members of the family compensate by verbally explaining to a blind person or by signing to a deaf person).

7.3 Multimodality as compensation for impairments

Multimodality, understood as the optional presentation of the same information content in more than one sensory mode, can compensate to a certain degree for sensory impairments.

Television and cinema shows are by nature multimodal events, but in most cases different modalities are used to convey different information, requiring seeing and hearing for fully understanding and appreciating the plot. The principles of multimodality can be used to compensate the lack of one sensory mode in a person by presenting the information in another mode. For television, this can be done by subtitling (captioning), i.e. a textual representation of the dialogue and the sounds (the representation of dialogues, only as can be found in some broadcast and DVD programmes, is only of limited use to deaf people if reference to sound events like shots or thunder is missing).

Similarly, not being able to see the action on the screen can be compensated by audio commentary which may possibly be somewhat disturbing to a seeing audience. Programmes broadcast with audio commentary for blind people (usually transmitted on the second stereo channel) are still rare, even rarer are audio commentary tracks for blind people on DVD (audio tracks are, however, often used for the director's comments).

The user interfaces of almost all devices and services can be designed or extended to encompass the principle of multimodality. An optional voice-prompt can lead a blind person through the process of withdrawing money from an ATM (security issues to be considered), visual menus can support the deaf person using a telephone e.g. for text messaging. In all cases, the redundant modalities have to be selected carefully, since e.g. many pre-lingual deaf people cannot read or write, just as many blind people cannot read Braille.

7.4 Alleviating social exclusion

The need for making an increasingly knowledge-based society accessible to everybody including disabled people is acknowledged by national and international governmental and regulatory institutions (see note). The insufficient presence of multimodality in our society is one of the factors contributing to the social exclusion of disabled people. Multimodality can close the gap for many disabled people, making the world and its opportunities more accessible to them and thus contribute to a reduction of the social exclusion of disabled people in their social and professional life. Introducing multimodality is a necessary, but by no means sufficient first step for overcoming social exclusion.

NOTE: Delivering Accessibility: Improving disabled people's assess to the knowledge based Society. Commission Staff Working Paper. Commission of the European Communities, Brussels. SEC(2002) 1039.

7.5 Benefits of multimodality for everybody

Multimodality has the potential of benefiting disabled as well as non-disabled people. Viewers watching a film in a language they do not understand appreciate subtitles in their own language. Many of those who are somewhat familiar with the language spoken in the film enjoy watching the film with subtitles in the original language (e.g. a Spanish viewer watching a French film with French subtitles for support).

Voice output of text information benefits seeing people in situations in which the visual channel is not available for reading, such as when driving a car.

8 Multimodality for improving accessibility

8.1 What is Accessibility?

Providing accessibility means removing barriers that prevent people with disabilities from participating in substantial life activities, including the use of services, products, and information. We see and use a multitude of access-related technologies in everyday life, many of which we may not recognize as disability related when we encounter them. Accessibility is by definition a category of usability: software that is not accessible to a particular user is not usable by that person. As with any usability measure, accessibility is necessarily defined relative to user task requirements and needs. For example, a telephone booth is accessible (e.g. usable) to a blind person, but may not be accessible to a person using a wheelchair. Graphical user interfaces are not very accessible to blind users, but relatively accessible to deaf users.

There is a distinction between "direct" access and access through add-on assistive technologies. Direct access can be described as "adaptations to product designs that can significantly increase their accessibility". A major advantage of this approach is that large numbers of users with mild to moderate disabilities can use systems without any modification.

Assistive access means that system infrastructure allows add-on assistive software to transparently provide specialized input and output capabilities. For example, screen readers allow blind users to navigate through applications, determine the state of controls, and read text via text to speech conversion. On-screen keyboards replace physical keyboards, and head-mounted pointers replace mice. These are only a few of the assistive technologies users may add on to their systems.

8.2 Assistive technologies

In this clause are discussed some of the needs, capabilities, and assistive technologies used by people with disabilities. The brief descriptions in clause 8 do not constitute complete coverage of the wide range of disabilities, capabilities, needs, and individual differences across the population of people with disabilities - we have focused on providing a broad introduction to visual, hearing, and physical disabilities. Users with cognitive, language, and other disabilities may have needs in addition to those discussed in this clause.

Use of assistive technologies varies across users and tasks. The discussion of assistive technologies is not comprehensive, but it does cover many commonly used software and hardware solutions. In reading clause 8, it is important to remember that as with all users, the needs of users with disabilities vary significantly from person to person. Many users with disabilities do not use assistive technologies, but can benefit from small design changes. Other users have significant investments in assistive technologies, but they too can benefit from software that better responds to their interaction needs.

8.3 Physical disabilities

Physical disabilities can be the result of congenital conditions, accidents, or excessive muscular strain. By the term "physical disability" we are referring to disabilities that affect the ability to move, manipulate objects, and interact with the physical world. Examples include spinal cord injuries, degenerative nerve diseases, stroke, and missing limbs. Repetitive stress injuries can result in physical disabilities, but because these injuries have a common root cause, we address that topic below under its own heading.

Many users with physical disabilities use computer systems without add-on assistive technologies. These users can especially benefit from small changes in interface accessibility. An example is a person who uses a standard PC and off-the-shelf Windows® business productivity applications, but navigates almost exclusively via the keyboard because his poor fine motor coordination makes a mouse or trackball difficult to use. When a pointing device becomes necessary, he uses a trackball.

Some users with physical disabilities use assistive technologies to aid their interactions. Common hardware add-ons include alternative pointing devices such as head tracking systems and joysticks. The MouseKeys® keyboard enhancement available for MS Windows®, Macintosh®, and Windows®-based workstations allows users to move the mouse pointer by pressing keys on the numeric keypad, using other keys to substitute for mouse button presses. Because system-level alternatives are available, it is not necessary for applications to provide mouse substitutes of their own. The problem of the mouse is a good example of the kind of generic issue that must be addressed at the system rather than application level.

Unfortunately, the MouseKeys® feature is often time-consuming in comparison to keyboard accelerators, because it provides relatively crude directional control. For tasks requiring drag and drop or continuous movement (e.g. drawing), MouseKeys® is also inefficient. On the other hand, because current systems are designed with the implicit assumption that the user has a mouse or equivalent pointing device, many tasks require selecting an object or pressing a control for which there is no keyboard alternative. In these cases, MouseKeys® provides an option. It is clear that future operating environments need to offer effective alternatives for users who may not use a pointing device.

It is important that applications provide keyboard access to controls, features, and information in environments that have keyboard navigation. Comprehensive keyboard access helps users who cannot use a mouse. Many environments allow users to use tab and arrow keys to navigate among controls in a window, space bar and enter to activate controls, and key combinations to move focus across windows. In some cases, extra engineering may be required to ensure that these features work in all areas of an interface.

In addition to keyboard navigation, keyboard accelerators and mnemonics are also helpful for users with physical disabilities (as well as blind and low vision users). Whenever practical, commonly used actions and application dialogs should be accessible through buttons or keyboard accelerators. Unfortunately few of the standard accelerator sequences were designed with disabilities in mind. Many key combinations are difficult for users with limited dexterity nonetheless, use of key mapping consistent with guidelines for the local application environment not only speeds use of applications for users with movement difficulties, but it also increases the effectiveness of alternate input technologies such as speech recognition. Assistive technologies often allow users to define macro sequences to accelerate common tasks. The more keyboard access an application provides, the greater the user's ability to customize assistive technology to work with that application.

Table 1: Assistive Technologies (AT) for physical disabilities

Assistive Technology	Function provided
Alternate Pointing Device	Gives users with limited or no arm and hand fine motor control the ability to control mouse movements and functions. Examples include foot operated mice, head-mounted pointing devices and eye-tracking systems.
Screen Keyboard	On-screen keyboard which provides the keys and functions of a physical keyboard. On-screen keyboards are typically used in conjunction with alternate pointing devices.
Predictive Dictionary	Predictive dictionaries speed typing by predicting words as the user types them, and offering those words in a list for the user to choose.
Speech Recognition	Allows the user with limited or no arm and hand fine motor control to input text and/or control the user interface via speech.

Table 2: Keyboard enhancements

Feature	Function provided
StickyKeys	Provides looking or latching of modifier keys (e.g. Shift, Control) so that they can be sued without simultaneously pressing the keys. This allows single finger operation of multiple key combinations.
MouseKeys	An alternative to the mouse which provides keyboard control of cursor movement and mouse button functions.
RepeatKeys	Delays the onset of key repeat, allowing users with limited coordination time to release keys.
SlowKeys	Requires a key to be held down for a set period before keypress acceptance. This prevents users with limited coordination from accidentally pressing keys.
BounceKeys	Requires a delay between keystrokes before accepting the next key press so users with tremors can prevent the system from accepting inadvertent key presses.
ToggleKeys	Indicates locking key state with a tone when pressed, e.g. Caps Lock.

8.4 Low vision

A related problem for users with low vision is their limited field of view. Because they must use large fonts or magnify the screen through hardware or software, a smaller amount of information is visible at one time. Some users have tunnel vision that restricts their view to a small portion of the screen, while others require magnification at levels that pushes much of an interface off-screen.

A limited field of view means that these users easily lose context. Events in an interface outside of their field of view may go unnoticed. These limitations in field of view imply that physical proximity of actions and consequences is especially important to users with low vision. In addition, providing redundant audio cues (or the option of audio) can notify users about new information or state changes. In the future, operating environments should allow users to quickly navigate to regions where new information is posted.

Interpreting information that depends on colour (e.g. red = stop, green = go) can be difficult for people with colour vision impairments. A significant number of men are unable to distinguish among some or any colours. As one legally blind user who had full vision as a child told us, his vision is like "watching black and white TV". In any case, a significant portion of any population will be "colour blind". For these reasons, colour should never be used as the only source of information - colour should provide information that is redundant to text, textures, symbols and other information.

Some combinations of background and text colours can result in text that is difficult to read for users with visual impairments. Again, the key is to provide both redundancy and choice. Users should also be given the ability to override default colours, so they can choose the colours that work best for them.

8.5 Blindness

There is no clear demarcation between low vision and true blindness, but for our purposes, a blind person can be considered to be anybody who does not use a visual display at all. These are users who read Braille displays or listen to speech output to get information from their systems.

Screen reader software provides access to graphical user interfaces by providing navigation as well as a Braille display or speech-synthesized reading of controls, text, and icons. The blind user typically uses tab and arrow controls to move through menus, buttons, icons, text areas, and other parts of the graphic interface. As the input focus moves, the screen reader provides Braille, speech, or non-speech audio feedback to indicate the user's position. For example, when focus moves to a button, the user might hear the words "button - Search", or when focus moves to a text input region, the user might hear a typewriter sound. Some screen readers provide this kind of information only in audio form, while others provide a Braille display (a series of pins that raise and lower dynamically to form a row of Braille characters).

Blind users rarely use a pointing device, and as discussed above, typically depend on keyboard navigation. A problem of concern to blind users is the growing use of graphics and windowing systems. The transition to window-based systems is an emotional issue, evoking complaints from blind users who feel they are being forced to use an environment that is not well suited to their style of interaction.

Although blind users have screen reading software that can read the text contents of buttons, menus, and other control areas, screen readers cannot read the contents of an icon or image. In the future, systems should be designed that provide descriptive information for all non-text objects. Until the appropriate infrastructure for providing this information becomes available, there are some measures that may help blind users access this information. Software engineers should give meaningful names for user interface objects in their code. Meaningful names can allow some screen reading software to provide useful information to users with visual impairments. Rather than naming an eraser graphic "widget5", for example, the code should call it "eraser" or some other descriptive name that users will understand if spoken by a screen reader.

Without such descriptive information, blind or low vision users may find it difficult or impossible to interpret unlabeled, graphically labelled, or custom interface objects. Providing descriptive information may provide the only means for access in these cases. As an added selling point to developers, meaningful widget names make for code that is easier to document and debug.

In addition to problems reading icons, blind users may have trouble reading areas of text that are not navigable via standard keyboard features. In OpenWindows® and MS Windows®.

EXAMPLE: It is not possible to move the focus to footer messages. If this capability were built into the design, then blind users could easily navigate to footer messages in any application and have their screen reading software read the content.

Table 3: Assistive Technologies (AT) for low vision and blind users

Assistive Technology	Function provided
Screen Reader Software	Allows users to navigate through windows, menus, and controls while receiving text and limited graphics information through speech output or Braille display.
Braille Display	Provides line-by-line Braille display of on-screen text using a series of pins to form Braille symbols that are constantly updated as the user navigates through the interface.
Text to Speech	Translates electronic text into speech via a speech synthesizer.
Screen Magnification	Provides magnification of a portion or all of a screen, including graphics and windows as well as text. Allows users to track position of the input focus.

8.6 About hearing disabilities

People with hearing impairments find it difficult to detect sound and may thus have difficulty distinguishing relevant audio information from typical background noise. Because current user interfaces rely heavily on visual presentation, users with hearing related disabilities rarely have serious problems interacting with software. In fact, most users with hearing disabilities can use off-the-shelf computers and software. This situation may change as computers, telephones, and video become more integrated. As more systems are developed for multimedia, desktop videoconferencing, and telephone functions designers will have to give greater consideration to the needs of users with hearing impairments.

Interfaces should not depend on the assumption that users can hear an auditory notice. In addition to users who are deaf, users sitting in airplanes, in noisy offices, or in public places where sound must be turned off also need the visual notification. Additionally, some users can only hear audible cues at certain frequencies or volumes. Volume and frequency of audio feedback should be easily configurable by the user.

Sounds unaccompanied by visual notification, such as a beep indicating that a print job is complete, are of no value to users with hearing impairments or others who are not using sound. While such sounds can be valuable, designers should not assume that sounds will be heard. Sound should be redundant to other sources of information. On the other hand, for the print example above, it would be intrusive for most users to see a highly visible notification sign whenever a printout is ready. Visual notices can include changing an icon, posting a message in an information area, or providing a message window as appropriate.

Again, the key point here is to provide users with options and redundant information. Everybody using a system in a public area benefits from the option of choosing whether to see or hear notices. When appropriate, redundant visual and audio notification gives the information that is necessary to those who need it. If visual notification does not make sense as a redundant or default behaviour, then it can be provided as an option.

Other issues to consider include the fact that many people who are born deaf learn American Sign Language as their first language, and English as their second language. For this reason, these users will have many of the same problems with text information as any other user for whom English is a second language, making simple and clear labelling especially important.

Currently, voice input is an option rather than an effective general means of interaction. As voice input becomes a more common method of interacting with systems, designers should remember that many deaf people have trouble speaking distinctly, and may not be able to use voice input reliably. Like the other methods of input already discussed, speech should not be the only way of interacting with a system.

Table 4: Assistive Technologies (AT) for hearing disabilities

Assistive Technology	Function provided
Telecommunications Device for the Deaf (TDD)	Provides a means for users to communicate over telephone lines using text terminals.
Closed Captioning	Provides text translation of spoken material on video media. Important computer applications include distance learning, CD-ROM, video teleconferencing, and other forms of interactive video.
ShowSounds	Proposed standard would provide visual translation of sound information. Non-speech audio such as system beeps would be presented via screen flashing or similar methods. Video and still images would be described through closed captions or related technologies. This capability would be provided by the system infrastructure.

8.7 Communication and language disabilities

Multimodal interaction presents some fundamental problems for users with impaired speech and language abilities. The use of a speech based interface for users with impaired speech will need to be tuned to the specific speaking abilities of these users. Speech recognition as an interaction modality is often seen as an important alternative to typed text as many of these users have additional dexterity impairments. In these cases however, the recognition dictionaries will need to be specifically tuned to the reduced set of sounds that the users can usefully produce.

Where users have reduced language abilities symbols based representation systems are generally used as an alternative communication medium. Work is currently underway [WAACC 2002] to develop a translation system between symbol strings or sequences into grammatically correct text strings using a generic content coding scheme to hold the content in a delivery independent manner.

Table 5: Assistive Technologies (AT) for users with impaired communication and language abilities

Assistive Technology	Function provided
Symbol Interface	Allows users to interact and communicate using symbols rather than text.
Reduced language complexity	Provides users with an alternative representation of content that has less complex vocabulary and language forms.
Speech Recognition	Can provide a working alternative to typed text entry provided the recognition dictionaries are tuned to the reduced range of sounds produced by some of these users.

9 Human factors guidelines and recommendations for multimodal interaction

This clause summarizes human factors guidelines for the design of systems with multimodal interaction. These guidelines describe various aspects of multimodal interfaces. They may be used to get answers to the question listed in this clause:

- In which application areas can the rules be applied, e.g. e-government, e-work, e-education, e-commerce, Teleshopping, e-leisure, e-banking, e-care, e-health).
- What are the representational building blocks used to create multimodal user interfaces?
- What are the transactional building blocks used to create complete transactions, e.g. authorization, data-entry, confirmation, searching, data presentation?
- Which actions/dialog components have to be considered in the design of multimodal user interfaces: e.g. conditions, navigation, dialogue control, repair, undo, recovery, selection of specific modalities for the presentation of an object?
- How do the location of the device used and the environment of the end-user device influence the selection of devices and modalities?
- How do abilities of different users affect the selection of modalities, where and when is the selection done?
- What are the effects of aesthetics, social acceptability, life style, pleasurenomics, branding and product differentiation on multi-modal user-interfaces?
- What needs to be done to create an atmosphere encouraging to complete the transaction in a multimodal interface?
- Which rules are implied by legal issues, e.g. on the accuracy of conveying information during a transaction?
- Where will the selection of the presentation style of contents or of the contents itself be done, at the data source, in the communication network or at the end-user device?

9.1 Metaprinciples for the design of multimodal systems

Metaprinciple 1: Use multimodal presentation of information to allow users with different preferences and abilities to interpret data in their preferred way

Description: Users have different abilities and special needs as well as individual ways of interpreting data. While some users may get maximum information from tables, others will prefer graphics or a combination of both; concurrently blind users may not be able to make any sense from either of these representations and need a spoken meta-description of the data in order to understand the relevant content.

Metaprinciple 2: Use multimodal interaction to allow users to interact with a system following their individual preferences and suited to their special needs

Description: Multimodal interaction is a necessary prerequisite for users with different preferences and abilities to interact with a uniform system. Input media allowing blind, deaf or motor impaired people to interact must be usable in conjunction, either sequentially or in parallel, as well as individually following the preference of the system users.

From these two meta-principles we can infer a number of important principles for the design of multimodal systems derived from usage requirements of system users.

9.2 Design principles for multimodal systems

The principles described in this clause are derived from the meta-principles described above. If applied properly, they will support system designers in creating truly usable multimodal systems. They are not intended to replace other, existing design rules which have been developed by various organizations to support the design of multimodal interfaces in specific environment settings. Consideration should be given to the adoption of these existing accessibility guidelines (e.g. W3C Accessibility Initiative) where appropriate.

Principle 1: The range of available modalities should be offered in the different modalities

Description: If a system offers different interaction modalities they need to be presented in all available presentation modalities for different users to be able to recognize their options. Otherwise people may not become aware of the options they can chose from.

EXAMPLE 1: If a system supports Braille output users should get this information both on the display and with auditory feedback, including advice on where the Braille output device is located. This allows blind users to realize that they may get information on the Braille output without seeing the device or being forced to search the output device.

EXAMPLE 2: Availability of features supporting deaf users must be presented in visible form (e.g. availability of close captioning, and the way to invoke such a feature (presentation of the videotext page, how to reach the relevant page).

EXAMPLE 3: The option for orally presented descriptions of movie storylines must be presented in an audible form, again including the description on how to invoke the service.

Principle 2: It should be possible to choose different presentation modalities using any of the available interaction modalities

Description: The user must be able to choose between different modalities using the interaction mode that he is comfortable with or able to use.

EXAMPLE: Switching between audio output and Braille output should be possible using either keypad input or a speech recognition system interpreting spoken commands.

Principle 3: Individual modalities should be activated optionally

Description: Users must have the option to not use specific modalities as well as the option to choose their preferred interaction and presentation modality.

EXAMPLE: Blind users may prefer spoken in- and output while being in a closed environment where they cannot be overheard. In public environments they probably prefer alternative modalities, which protect their privacy (e.g. Braille output).

Principle 4: Individual modalities should be scalable by the user

Description: Features of individual modalities (like the audio level, display contrast etc.) need to be adjustable to both environmental conditions and individual abilities and preferences.

EXAMPLE: Audio output must be adjustable to background noise; brightness of a display may be automatically adjusted to environmental conditions, force feedback should be adjustable to individual strength.

Principle 5: Completeness of functionality/controls should be ensured in all available modalities

Description: all functions offered by a system to the user must be addressable with all offered modalities. These functions include the functionality used to control the flow of action and the dialog between user and system: undo-facilities, error-recovery, redo-mechanisms, interrupt facilities etc.

EXAMPLE: A blind user must be able to undo errors during input to a system. If the undo-button is only visible on the screen and there is no keyboard input or shortcut to undo-functionality, a blind user will not be able to perform the desired undo-function. It is therefore required that undo and other related functionality is accessible by Principle 6.

Principle 6: Modalities should only be switched at logically sensible situations in the dialogue

Description: There is a smallest logical unit within which switching of modalities is not a sensible user action. Within these logical units no modality switch should take place, neither on the output nor on the input side. It is foreseen that these smallest logical units will be a complete transaction step.

EXAMPLE: Input of a name when doing online-booking should be done either by spelling to a voice recognition system, typing on a keyboard or on a touch pad or some other appropriate input means. It is not, however, necessary to offer the possibility to switch between these means deliberately. Note that correcting false input by other input means is perfectly acceptable, since correction of false input is a new step in the transaction that the user intends to perform.

Principle 7: The user-specific modality setting should persist

Description: The setting of modalities which is optimized by the users for their specific preferences and/or abilities should not be changed without explicit request from the users. This persistence of modality setting has two distinguishable aspects: (i) during one session the system behaves consistently, and (ii) a system behaves consistently over an extended period of time, possibly even across access to a system using different devices. If a system is designed to support such a persistence across devices, locations and time, it either has to integrate a clear user model (accessible from each device and password-protected, or the user has the information about his preferences with him, in the form of a smart card or an access device which can connect to the system at any given access point.

EXAMPLE 1: A use which has set a system interface for audio output at high volume in a noise environment in which he may not be able to use touch input (e.g. a car) wants his setting to remain while he is in this environment. Neither a change of output volume nor being forced to use his hands for input would be acceptable to him. (Note that it might be possible in certain circumstances that the system modifies certain setting dependant on the environment (e.g. speed-dependant audio-volume-setting in a car).

EXAMPLE 2: Users with certain disabilities preventing them from using specific input or output channels (e.g. blind people unable to use touch screens) do not want to set up an automatic teller machine to alternate modality settings every time they want to draw money from their bank. Their required modality setting should be "remembered" by the system, possibly stored with the card information or the account number and set each time the user accesses the system any access point.

Principle 8: Adequate prompts should be offered in the chosen modality

Description: If a user cannot use a specific output channel (e.g. deaf or blind users) the system must offer prompts (i.e. input requests) in alternate modalities accessible to the user. This principle also covers the use of so-called softkeys. If a softkey is being labelled with text in a graphical representation, then an alternate audio description must be available for people with limited eyesight.

EXAMPLE: Blind users often have problems using their mobile phones if the user interface of the device makes have use of graphically represented softkeys. A multimodal interface should support these users by either the possibility to get audio-descriptions of the actual key functionality or by modifying the system in a way that there are alternative ways to access the functionality "hidden" behind theses softkeys. Alternate possibilities would be the representation of softkey-labels in Braille-format or using some form of vibration-code to supply the information. Where audio dialogues are provided as an alternative they need to let people know the current context or mode- e.g. provide "where am I" option.

Principle 9: The same information should be expressed in different modalities

Description: The information supplied by a transaction system should be the same in all modalities wherever possible. It is very important that all essential information and the sense and emphasis of the information are the same in all modalities. This principle should, however, not prevent system designers to integrate rich information representation, which cannot be transposed in other representations, where adequate. If this case occurs users unable to access the information should at least get the information that user might be able to get more information.

EXAMPLE 1: In a hotel reservation system users should get the option to access all relevant information about chosen hotels in graphical, written and auditory form. Even if the system supplies additional information not accessible to blind people (e.g. a video of the hotel room or of the view over the seaside from the window) the system should explain to blind users that the video is available and describe the contents so they know about the possibility to get more information through human help.

EXAMPLE 2: In a TV system offering close captions for deaf users the information in the caption and the audible dialog should always carry the same information. If this is not possible due to the speed of the dialogue relevant information must be transmitted to have the user understand the dynamics of the story. For blind people audio descriptions of movie scenes delivered in parallel through a radio channel may be helpful to follow video transmission. Even if time does not permit to describe scenes in detail, relevant information ("the gardener carrying a rose bush is entering from the greenhouse") should be available.

Principle 10: The same style should be expressed in different modalities

Description: The interaction style as well as the style of content representation in multimodal systems should be consistent across different modalities. Transactional systems are being built for users to achieve certain goals and should support these activities through interfaces making the experience fun and the system a pleasure to use.

Designers know about the importance of aesthetics and fun for the success of a system. These factors should be carried across to other modalities if a user is unable to use a specific modality for the interaction with the system.

EXAMPLE: The use of flashing elements on a screen to attract the attention of a user should be presented in audible form, too. Additionally, a user with limited eyesight may need information on how to access the functionality connected to this "flashing" information. "Click here to get more information" would not be very helpful for a blind user.

Principle 11: Dialogues should be aesthetically consistent in each and every modality

Description: Dialogs in multimodal transaction systems should be designed in a way that they give a consistent aesthetic picture to the user across the transaction. This does not exclude that a dialogue is designed in alternative presentations for different user groups (i.e. for children, adolescent users or older people), it is, however, relevant that there is no disrupting change in the aesthetics of presentation of dialog control which would disturb the user during his effort to achieve a certain goal. This consistency of aesthetics in representation should carry over different modalities.

EXAMPLE: The user should not have to learn a new user interface and style of interaction when dealing with money transactions through a mobile (PDA) interface after having experienced a pleasurable user interface either on his personal computer at home or at a teller machine in his bank. The quality of the interface should be consistent over there different devices (offering different input and output modalities) and over the entire time of the transaction in question. Also, given that a banking application is to be used for different kinds of transaction it is helpful to the user if the aesthetic quality of the interface is consistent across these transactions.

Principle 12: Multimodality should not be stigmatizing

Description: Caution should be exercised to ensure that multimodality is not used in an insensitive or stigmatizing way. This includes a certain restriction which designers should accept: never design a system that cannot be used by people lacking the ability to use one specific modality. Inability to use a system at all has a clearly stigmatizing effect.

EXAMPLE 1: In the wording of a system's user interfaces it is important not to use any vocabulary that may be offending to a specific user group. It is important that, although some people may have specific requirements for interaction, it must be obvious in the interface that they are not valued by their specific needs. Also, care should be taken that the industrial design of interface devices does not stigmatize specific user groups.

EXAMPLE 2: A system using graphical representation should never use "moving buttons", i.e. access to functionality which cannot be used by people with motor limitations.

9.3 Implementation principles

Principle 13: Content should be adequately designed to allow for consistent multimodal presentation

Description: Content should be stored in delivery-independent form to allow for translation into the diversity of modalities. The translation into different representation should be done as far "down the line" as possible.

EXAMPLE: It is almost impossible to create a usable audio interface to represent data which was originally stored in a way suitable for graphical presentation. If a multimodal system is to represent contents in such diverse modalities it is a prerequisite that the data organization of the contents does not reflect specific properties of one of these modalities. The decision on how contents are to be presented should be left to the user (see principle 3). To allow for maximum adaptation to users' preferences it is advisable to transpose modality-independent data into modality-dependant presentations at the users' site.

Principle 14: A system should be designed to allow for consistent multimodal interaction

Description: Translation of input events of one modality into events in another modality to fit the requirements of a system that has been designed specifically for this second modality is, as a rule, not possible. The system design should provide a modality-independent description of dialogues and the respective control facilities.

EXAMPLE: While audio-driven dialogues tend to have a fairly clear structure and very little obvious parallelism in the user interface, point and click interfaces allow for maximum parallelism and many options to the user. If a system is to allow for both styles of interaction it is a prerequisite that the dialog is represented in a way that the system can translate dialog actions optimized to one user's preferences at the latest possible point in time.

10 Conclusions and recommendations for future work

The results in the present document summarize important principles applicable to the design of multimodal systems. These principles, while describing important design issues, are on a very generic level, and in-depth design expertise is needed to apply them to real world problems. An important next step should be to try to operationalize these principles into a set of design guideline, which can be easily applied by any experienced user-interface-designer during his design activities.

Experienced usability experts can develop these operational design rules without further in-depth studies of user behaviour, it is, however, recommended to apply the resulting rules in the design of pilot systems before putting them into widespread use. At least some of these guidelines may only be applicable to or targeted to specific application areas like e.g. teleshopping or tele-administration.

The greatest opportunity for commercial service delivery using multimodal interfaces may be in the area of "independent living", because there a high demand will develop due to the changing age spectrum in our civilization. It is therefore our suggestion to put priority on services for "independent living" when starting to develop design guidelines for user interface designers of multimodal transactional systems using delivery-independent contents representations.

On the implementation level low-cost adaptation at the design phase of service deployment rather than extensive retrospective adaptation at the delivery phase is the preferred method of design. Modality-independent storage of content seems to us the implementation of choice compared to translation between different modalities at presentation time. The feasibility of such an approach has been shown by various systems. It seems to us less the question of designing new architectures than doing a well-founded selection from different, already existing architectures.

The convergence of different ICT areas (internet, telecom, computing) may be an important reason for ETSI to invest in new areas like architectures for content storage, applications or service delivery, or at least to coordinate with other relevant standardization bodies working in these fields.

Other issues that have only been scratched on the surface by the present document deal with questions about how the diversity of cultures, life-style or uptake of technology advances may be influenced by multimodal interfaces in a positive way. It is plausible that technologies used for the implementation of multimodal systems can be used to ease further integration of European cultures through technology advances. This topic certainly deserves further research.

Similarly, multimodal systems may not only be used to close the gap between disabled and fully-abled people, but also between generations, different cultures or different languages. The weighting of the design principles in systems trying to solve these issues may be very much culture-dependant; these dependencies are not very well understood and require further investigation.

Similar questions of evaluating the importance of design principles arise if different user groups become the target of the design of transactional systems. While the validity of the design principles remains, the importance of their application may change if systems are designed for children, teens, young adults, or the "average" user of today's systems. Also, new design rules may become relevant which could not be developed from the experience with people with special needs.

Other topics which deserve further investigation are the effects of nomadic and ubiquitous access and the requirement of being independent of computing equipment and communication systems properties on the design rules and their applicability. The rules in the present document are based on experiences with PCs and PDAs for the access to transactional systems over the Internet. If these requirements change the design principles inferred may also change and the relevance of individual rules may change. A detailed analysis of these effects was beyond the scope of the present document.

Finally, operational success of multimodal system is not all that counts. Even the most usable system will only have the success if its users enjoy working with the system and deliberately come back to use it again. Aesthetics and desirability, important means to reach the goal of creating truly useful systems need to be taken into account during system design. How this can be done and how these factors influence the design of systems deserves further investigation.

Annex A (informative): Stakeholder consultations

A.1 Questionnaire(s)

ETSI

European Telecommunications Standards Institute

Introduction

This is a questionnaire issued by ETSI, the European Telecommunications Standards Institute which develops international standards for telecommunications networks, terminals and services. One of ETSI's major concerns is to make sure that its standards take into account the needs of the widest possible range of users including elderly users and users with disabilities.

The questionnaire addresses the fact that users who suffer from an impairment of one or possibly more sensory abilities, e.g. visually impaired or blind users, cannot make use of a device or service, if the information needed for using the device or service is presented in one sensory modality only. For instance, a blind user cannot make use of information presented only visually on a display; a deaf person cannot make use of a voice-menu based telephone service.

In order to improve the usability to tomorrow's consumer devices and services, ETSI is preparing a set of guidelines on how to present information in a multimodal way. Multimodal user interfaces offer to the user the choice of different input and output modalities (e.g. acoustic, visual or tactile) to compensate for impaired sensory abilities. This is achieved by the automatic transformation of the content from one modality (such as text) to another modality (such as synthetic voice).

To make sure that the users' needs are well captured and that possible existing and future solutions are taken into account, we have designed this questionnaire which focuses in particular on telecommunications devices and transactional services such as internet banking, teleshopping web-sites and voice-menu based telephone services.

ETSI Secretariat

Postal address: F-06921 Sophia Antipolis Cedex - FRANCE

Office address: 650 Route des Lucioles - Sophia Antipolis - Valbonne - FRANCE

Internet: secretariat@etsi.fr

Tel.: +33 4 92 94 42 00 - Fax: +33 4 93 65 47 16

A) Background about yourself

In this first section of the questionnaire, we would like to ask you to provide some information about yourself and the organization you represent.

A.1) Please indicate who you are and how you can be contacted:

Name: ______________________________

Address: ______________________________

____

____

Phone: ______________________________

Email: ______________________________

A.2) What is your role in the organization that you represent?

____

____

A.3) Which of the following describes best the type of organization you represent? (select one of the following):

Representation of visually-impaired people

Representation of deaf and hard-of-hearing people

Representation of mobility-impaired people

Representation of elderly people

Other representation of users

Other non-government organization (NGO)

none of the above, please specify:

____

B) Background about the organization

B.1) Which type(s) of disability are the members of your organization faced with?

____

____

____

B.2) How many members does your organization have (number of disabled members only)?

____

____

____

B.3) How many people in your country are faced with the disability/disabilities you mentioned above?

____

____

B.4) Do you run panels/discussion groups with your members on problems they face due to their disability when using technical products (telephones, computers, TVs, etc.)?

no

yes
(if yes, what kind of panes/discussion group?)

____

____

____

B.5) Do you give advice to your members about which technical products are particularly well suited for their requirements?

no

yes
(if yes, in what form do you give advice?)

____

____

C) Using transactional services

C.1) In the experience of your organization, are you aware of elderly or disabled users having difficulties using transactional services such as internet banking, teleshopping web-sites and voice-menu based telephone services?

no

yes
(if yes, which users?)

____

____

____

(if yes, what kind of difficulties?)

C.2) What solutions to these difficulties have you encountered?

C.3) Have you produced or are you developing any solutions to these difficulties?

C.4) What is your solution based on?
(select one of the following):

"in-house" experience

standardized methods

academic research

none of the above, please specify:

C.5) Is your solution available for use?

no

yes
(if yes, please give details)

C.6) Can the technique/technology in your solution be used by others, and under what purchase/licence arrangements?

no

yes
(if yes, please give details)

C.7) What further research needs to be done to ensure that transactional services can be made usable for older people and people with disabilities?

D) Using telephones

D.1) In the experience of your organization, are you aware of elderly or disabled users having difficulties using telephones (mobile or fixed-network phones) for voice or text conversations?

no

yes
(if yes, which users?)

(if yes, what kind of difficulties?)

D.2) What solutions to these difficulties have you encountered?

D.3) Have you produced or are you developing any solutions to these difficulties?

____

D.4) What is your solution based on?
(select one of the following):

"in-house" experience

standardized methods

academic research

none of the above, please specify:

D.5) Is your solution available for use?

no

yes
(if yes, please give details)

D.6) Can the technique/technology in your solution be used by others, and under what purchase/licence arrangements?

no

yes
(if yes, please give details)

D.7) What further research needs to be done to ensure that transactional services can be made usable for older people and people with disabilities?

D.8) What are the most important things the manufacturers of **mobile phones** should change in order to make their products usable for your members? Please list in order of importance.

1) ______________________________
2) ______________________________
3) ______________________________
4) ______________________________
5) ______________________________
6) ______________________________
7) ______________________________
8) ______________________________
9) ______________________________
10) ______________________________

D.9) What are the most important things the manufacturers of fixed-network phones (i.e. phones people use at home or at work) should change in order to make their products usable for your members? Please list in order of importance.

1) ______________________________
2) ______________________________
3) ______________________________
4) ______________________________
5) ______________________________
6) ______________________________
7) ______________________________
8) ______________________________
9) ______________________________
10) ______________________________

D.10) What are the most important things the service providers (e.g. network operators, directory enquiry services, call centres, etc.) should change in order to make their services usable for your members? Please list in order of importance.

1) ______________________________
2) ______________________________
3) ______________________________
4) ______________________________
5) ______________________________
6) ______________________________
7) ______________________________

8) ____

9) ____

10) ____

Thank your very much for your help. After filling in the questionnaire, please send it to [to be specified].

If you have any questions or comments, please contact:

[name]

[phone]

[Email]

A.2 Records of consultations

A.2.1 User Consultation, Perth, 20th June 2002

Session 1:

Run by: Nick Hine

Running: Nokia 8310 Mobile Phone, using WAP browser

Task: The task was to participate in a World Cup Football Quiz being run by the Orange mobile phone operator in the UK. This quiz was presented as a GPRS delivered WAP service.

The test was conducted as follows:

With an impaired user, the tester went through the whole procedure first, then handed the phone to the user for the user to attempt to use the phone to answer the quiz questions. This involved using the toggle and connection keys to move within menus and to select choices.

Script:

1) Navigate to **services** using the Menu and up down arrows

2) Select **home** with the select key

3) Select **fun & games** with the green key

4) Select **quizzes** with the green key

5) Select **World Cup Quiz** with the green key

6) Answer 3 questions with the green key

User: Da

D has moderately poor manual dexterity and poor speech as a result of cerebral palsy. She had difficulty using the small keys on the phone, and the phone had to be held for her. She appreciated the haptic feedback in the form of the positive click felt when the keys were pressed. She had no difficulty in following the task procedure, and enjoyed attempting the quiz.

User: K

K has muscular dystrophy and uses a powered wheelchair. He was able to hold the phone and use the required keys. He found the display small as he rested his hands on his lap while using the phone. He enjoyed attempting the quiz and had no difficulty following the required procedure.

User: G

G has fairly severely impaired manual dexterity and extremely poor speech as a result of cerebral palsy. He was unable to hold the phone and had considerable difficulty selecting the keys on the phone if it was held for him. He had no difficulty following the sequence of operations for the quiz, and he enjoyed attempting the quiz.

User: B

B has fairly uncontrolled movements and poor speech as a result of cerebral palsy. He was unable to hold the phone and press the buttons. When the phone was held for him, he had great difficulty in selecting keys. He had no difficulty following the sequence of operations for the quiz, and he enjoyed attempting the quiz.

User: R

Ross has a paralysed left side and impaired speech as a result of a car accident. He also has short-term memory difficulties. Prior to the accident he was training as a professional footballer and was interested to attempt the quiz. He was unable to hold the phone, but could select the appropriate buttons if the phone was held for him. The sequence of operations for the quiz were sufficiently simple for him to remember them, so he had little difficult with the quiz.

Key Issues: The small size of the physical interface presented some users with reduced dexterity with some significant problems. The positive tactile feedback was, however, considered to be valuable. All users had little difficulty following the sequence of operations for the quiz.

Opportunities: All the participants enjoyed the challenge of the quiz and enjoyed the sense of inclusion that accessing a mainstream service produced.

Removing Barriers: For some users, inclusion could only be achieved by providing the same service on a more accessible device.

Session 2:

Run by: Sian Joel

Running: Fujitsu LifeBook B Series

Task: The purpose this task was to register, obtain and interact with the RadioTimes online virtual plant game.

The demonstrator initially showed the user how to complete the task. This began by opening the browser and Radiotimes.com/plant website. Then the registration form was completed and submitted. A question relating to gardening was then shown which once answered correctly, enables the plant to grow. After answering the initial question, the Virtual Plant game is then played. It is demonstrated how the virtual plant can be fed, watered and tended to within the Virtual Garden.

Once this demonstration was completed, the user's ability was accessed to see if they were able to complete the task or any stage of the task. Where the user is able to, the task is completed, where the user is unable to, they were asked to comment on the various stages of the task demonstrated to them. If the user completes the task, they were asked for their opinion of the site and the difficulties they had.

User: Du

Impairment: Du has severe physical restriction (wheelchair), no speech (used Delta-talker) and is unable to read. Du is able to communicate relatively quickly through his Delta - talker. However the range of vocabulary is limited due to the restricted number of words being represented through pictures on the Delta - Talker.

Key Issues: Du was unable to carry out the task but gave his opinion on the demonstration presented. The first issue that Du expressed was the inability to read which meant he could not complete the registration. He commented that this inability meant he could not use the Internet in general. He could however use software which did not require him to read, for example where pictures represented the words or tasks. In the case of the Virtual Plant where certain actions were represented as icons, e.g. watering can icon for watering the plant, Du felt he could be able to do this part of the software. He also commented that he would only be able to use that aspect of the software if the hardware was more suitable than the laptop presented. He listed keypads, tracker-balls and in particular joysticks as being more appropriate.

Opportunities: With the "right" hardware or integration with the user's existing hardware (e.g. Delta - Talk), the Virtual Plant website could be completed relatively successfully. Acknowledging Du's inability to read and therefore inability to register and answer the initial gardening question, the Virtual Garden aspect of the website could be achieved thoroughly as tasks are represented as pictures.

Removing Barriers: The biggest barrier to Du is text. Incapable of reading the instructions, registering or answering the questions, he would find this site inaccessible at the first hurdle. This barrier is very difficult to remove as there is a need for registration in order to play the game, win the competition and ultimately gain the prize money (where would RadioTimes know where to send the prize money to without a registration of address?). One possibility is for the text to be spoken, however, text still needs to be inputted into the website. Another possible solution would be for an able bodied representative to complete the registration and question and for Du to complete the Virtual Garden aspect.

User: Da

Impairment: Da has medium/severe physical impairment (wheelchair) and uses a LightWriter to communicate. She can use her LightWriter quickly.

Key Issues: Da found that she could not complete the Virtual Plant task using the hardware that was presented to her. She would be able to use the software if it was somehow integrated with her existing devices or the laptop was smaller and she could place it on her lap or on her chair (noting that her LightWriter takes up most of the room on her wheel chair tray). Da thought the website looked good and the size of lettering and the icons were easily understood.

Opportunities: Da would be able to use the Virtual Plant site if it was integrated into her LightWriter or if the hardware device was smaller e.g. a PDA.

Removing Barriers: In terms of Da there are no barriers within the software itself, but within the device used to display it.

User: R

Impairment: R had a car accident which paralysed him down his left side. He also has short term memory problems.

Key Issues: R was able to complete much of the site with some help on what he should do and which aspects to fill in. R has used computers previously and had used the internet. He had not used a laptop touch pad mouse and preferred using a "normal" mouse. He found he was a little slow at using the touch pad but with practice he would be able to use it. He also found that with his unfamiliarity with the laptop mouse, he would slip from the main Virtual Garden screen. That screen, written in Flash, could not be maximized to the full monitor size and hence when R mouse clicked out of the screen it would be on the wrong window. This confused R as the Garden screen disappeared and the other window came to the forefront.

The main issue with R was his short-term memory limitations. This would mean that although he could complete the site with some help today, he would not be able to repeat the process again tomorrow.

Opportunities: Perhaps the primary opportunity relates to R's short-term memory limitations. The local care staff suggested that if all transactional websites had the same or standardized registration this would help people like R. Repeatedly carrying out registration tasks would then move into the sphere of long term memory and allow R to complete online registrations without intervention. Another possibility could be to use very basic or "low tech" solutions that mirrored traditional registration forms. If there were no graphics or additional information other than the form itself, the registration would be comparable to the paper registration which R knows and remembers how to complete. The addition of graphics and superfluous information confused R.

Removing Barriers: Similarly to the opportunities arrived at, the idea of standardization could be the key to removing the barrier of not remembering/knowing how to complete online registrations. However, although registration could be standardized (or in a regular format), it would be very difficult to standardize all transactional websites. The content of each website would have an identity of its own, with corporate branding and differing facilities and products being available.

User: Dr

Impairment: sever cerebral palsy, very limited speech.

Key Issues: Dr could not complete the Virtual Plant website game and said he would find it difficult using the laptop. Because Dr does not use a Delta Talk machine and his speech is difficult to understand, it was not easy to evaluate his possible use of transactional services.

Conclusions and Recommendations:

It is difficult to surmise conclusions and recommendations that can apply to every person in every circumstance. It goes without saying that everyone is difficult and each of the people listed above had very different issues with the software. Du for example could not read, and recommendations could be made about making the text vocalised. R on the other hand had short-term memory problems and recommendations could be considered to standardize registration on transactional sites.

In general, however, it did not seem that the software itself was the main issue but moreover the actual hardware used. Da and Du both commented they would find it easier if the laptop was somehow "different". It was suggested that the laptop be smaller to fit beside the Delta talk, the software somehow incorporated the Delta talk or that instead of a mouse a joystick could be used.

In conclusion therefore, recommendations can be found for each person who addresses each individual need. In general however, the laptop itself proved to be the biggest obstacle. This was especially the case with the Virtual Plant website, where its concept was very basic and most of the users above said they could use if the tools for using it were available. Some people in this research (Du, Da and Dr) physically could not even use the device that was presented before them. Therefore conclusions made on the device are perhaps more generic when discussing usability. Issues concerning the software content, on the other hand, are more related to the area of accessibility and arguably more centred around individual preferences and needs.

Session 3:

Run by: Susie Schofield

Running: HP Jornada running Windows CE 2.11 and the AvantGo browser 3.3

The task was to find out the time of a flight from Edinburgh to Oslo on 25th June 2002 using a Palm Pilot connected to the web. The user was then required to find out the telephone number of the airline.

The test was conducted as follows:

With an impaired user, the tester went through the whole procedure first, and then went through the task line by line while the user performed the task. Help was given when the user got stuck.

For an employee at the centre, the guidance was limited as I wanted to see how intuitive the web page was to someone without the additional physical dexterity problems.

Task: The task involves attempting to find the options available to fly from Edinburgh to Oslo on 25th June 2002, then to find the phone number of the airline that is the first choice.

1) Open the AvantGo Browser
2) Select the **Expedia To Go** service
3) Select the **Flight Timetables**
4) Enter Edinburgh in the From box
5) Enter Oslo in the To box
6) Enter 25th June 2002 into the Departing box
7) Submit
8) Find the first airline suggested
9) Go back to the Expedia home page
10) Find, if available, the phone number of that airline

User: IW (worker at the centre)

Too much text. Icons could be used more. User would need to be able to read to use this site. For many users, it would need to be anchored but this would negate its portability. For some users, anchored on a wall would be better than anchored onto table. For many users the PDA would be too flimsy. Would need to be more robust and waterproof. The stylus will be too fiddly for many users - touch would be better.

Cost may be prohibitive.

User: G

Impairment: severe physical restriction (wheelchair), no speech (used Delta-talker)

G communicates using Delta-talker, and is quite fast and accurate with this form of communication. The keyboard of the Jornada was far too small for G to use. He was able to hold the stylus in his left hand and stab at the Jornada, but with no accuracy. He also found the position he needed the Jornada to view the screen was not the same as the ideal position for using it. G was worried the Jornada would interfere with his Delta-talker, but liked the idea of a larger PDA. He liked the idea of being able to access the wealth of the web, and wondered if something the size of a Tablet with touch screen could one day replace his Delta-talker.

User: K

Impairment: palsied (wheelchair). Enjoys using computers and mobile phone

K was very keen on the Jornada and liked its size as it was easy for him to hold in his left hand. However, due to hand shake, he found the area he needed to click too small. He clicked the magnifying glass expecting this to zoom, but instead this was the icon for the find function. K was then unsure how to get out of the find function. When he pressed the keyboard icon, the keyboard appeared then disappeared due probably to K pressing the icon down too long. He would have liked a "sticky finger" accessibility feature.

When K typed in Edinburgh he was disconcerted when the display said "inburgh". Due to the size of the PDA, the software designer has scrolled the word but there is no visual clue to this. K assumed he had lost the first two letters and proceeded to delete the word and type it in again. He then noticed the text-predictor which he used to select Edinburgh. He thought this a very good feature.

K had clicked the caps lock at some stage and the numbers had changed into symbols. When it came to putting in the date, there were no numbers on the keyboard so K decided to submit his request for today's date. A message came up "sorry, no flights matching" with no clue as to how to get back. On pressing the home page, K ended up at the Orange site due to selecting the wrong Home page.

K was able to take the stylus out and replace it himself. Although he found the task difficult, he felt with practice he would not find this an insurmountable difficulty. He found getting the flick lid up difficult as his fingers were coming round the top of the device. He did not have the dexterity to hold it in such a way as to keep the lid free. Overall, K was very impressed with the PDA and liked the idea of wireless connection to the web. He preferred being able to hold the device to it being fixed either to the table or to his chair.

User: Dr

Impairment: sever palsy, very limited speech

Dr was unable to use the PDA. He found the display too small and the stylus too fiddly. He did not see himself using a PDA.

User: A

Impairment: None (helper)

A found the software difficult to navigate. She was unable to see flight timetables as they needed to be scrolled down to and she was unaware of this. Having found the timetables, Alison was unclear how to navigate. She tried to highlight the link rather than click on it. She was not able to type details in as again she did not realize she needed to click in the box before using the keyboard. Like K, she was alarmed when the beginning of Edinburgh disappeared, and again like K, she could not find the numbers to change the date.

She found the text very small, and did not feel a PDA would be of much use to her in its current format.

Session 4:

Run By: Stephen Furner

Running: On-line Shopping dialogue

This session provided visitors with an opportunity to try out the TESCO on-line shopping services. The services were demonstrated to the stand visitors and items suggested by them were searched for. The use of this type of dialogue, and issues for people with impairments that may influence its accessibility, were discussed with the stand visitors.

Description of demonstration: The demonstration was of the interface offered on the public Internet by TESCO for making purchases on line. It was demonstrated using a laptop computer connected to a local area network provided by Dundee University for the stands in the hall testing.

The demonstration consisted of searching for a product, such as a packet of chocolate biscuits, and going through the shopping dialogue up to the point of purchase. Initially, the presenter chose an item to show the dialogue. Once visitors could see how it worked they suggested items to be selected for purchase.

The presenter discussed the dialogue with visitors using the STF 204 questionnaire as the starting point for the discussions.

The initial demonstration to visitors made use of the shopping service provided for non-disabled customers by TESCO. Once the basic system had been shown the dialogue that has been provided by TESCO for disabled customers was demonstrated.

Stand visitors: The people visiting the stand were centre clients and centre workers. The centre clients who visited the stand were wheel chair users. The centre clients had a range of motor control problems that made conventional keyboard operation a difficulty. One of the visitors used a chin switch. One visitor used a computer based communications aid.

The centre clients who visited the stand were competent intelligent individuals who where familiar with the use of information technology. They could be expected to be customers for this type of shopping service if it was more convenient than personal shopping at a supermarket or local shop. Given the mobility problems of the stand visitors, this type of shopping service could be expected to offer benefits in time saving if it were sufficiently easy to use.

Key issues raised by stand visitors:

1) Dialogue needs to work with tracker ball

2) Dialogue needs to be easy to use

3) Need to be able to get same level help available face to face in shops

4) What about social aspects of shopping?

5) Current dialogue is essentially text based consequently requires good skills in English language

6) Hitting small targets such as little tick boxes will be a problem for some people with motor impairments

7) Authentication needs to be quick and simple - it can take a long time to enter logon sequence. Specifically if new password is needed

8) Cognitive skills are required to understand the navigation and content of the information being presented. This is in addition to the basic issues of physical access to the input and display devices

Opportunities for removing barriers: An access dialogue is offered by this shopping system for disabled people, this is primarily targeted at the provision of access for people with visual disabilities. It assumes that the user has good literacy skills and is able to communicate with a computer by text. With the increasing bandwidth currently becoming available from broadband communications services, there are opportunities here for the increased use of graphics for people who do not have good language skills.

The use of alternatives to text could make use of Avatar based shopping assistance and, if a computer based transaction is not possible, then human intervention could be offered from an Internet based call centre. While an Avatar cannot replace all of the capabilities of human support, it may be able to provide some of the benefits of human support to improve the functional capabilities of the dialogue and improve the customers' experience.

For people using personal communicators it may be simpler to connect directly to the shopping systems, via an infra red or radio link to a private or public network access point or via a dial up mobile data services, rather than attempt to make use of an adapted conventional desktop PC. This may make shopping access easier if the engineering of the basic service is sufficiently robust and comprehensive to enable access by the Internet browsers used by disabled people from this terminal technology.

The accessibility of this type of transactional services appeared to be a combination of the terminal and the content. It did not appear to simply a question of access to the input or display from a PC terminal but dependent on an interaction between terminal capability and the design of the content of the shopping service.

User Consultation, Bad Tolz

Run By: Lutz Groh

This document reports the outcome of a one-day-mission-trip to the premises of the Generation Research Program (GRP) in Bad Tölz. This program was initiated by the Human Studies Centre (HWZ) of the Ludwig-Maximilian University (LMU) in Munich and is focusing on interdisciplinary research into the measurement and improvement of the health, quality of life, and functional capacity of the general population, with particular attention to that of elderly individuals. All activities take place in close cooperation with the regional medical and psychosocial institutions for the development of new concepts and methods for the improvement of the quality of life of elderly individuals. The regional government of Upper Bavaria supports the project.

The meeting took place on 1st May 2002. It was conducted in form of an interview with Mrs. Sophia Poulaki, geronto-psychologist and head of the Research-Group "Competence in old age" and Dr. Miklos Kiss, head of the project group "Man Machine interface". Both gave a short overview on the work of the GRP in general and on their projects in detail. Mrs Poulaki, whose main field of work is currently the competence in the so-called 4th age (75+), outlined that age and aging is influenced by many factors, and that elderly people have to cope with permanently changing options. Physical, cognitive and environmental changes are a challenge and often require developing new coping strategies. Since the elderly are a very inhomogeneous group and the process of aging is highly diverse and takes place over an extended period of time, there exist various physical, functional, and cognitive demands, as well as different living conditions and lifestyles. Persons between 60 years and 75 years old are said to fulfil the ideal of "late freedom", whereas people over 75 are perceptible for many physical and cognitive illnesses. Those people are increasingly being confronted with different psychological and social changes. That of course, does not mean that age has to be brought automatically in context with illness. Age-depending weaknesses in terms of physical and cognitive capabilities have either to be compensated by smart, easily comprehensible and specifically designed, technical means or, as it is done by Mrs. Poulaki's group "Competence in Old Age", by developing programs for the maintenance of competence and autonomy in old age.

Mrs. Poulaki stated that competence in this context means the ability to develop appropriate strategies to cope with the dissonance between prevailing requirements and capabilities to find the optimal solution for one's own interests and intentions.

For that reason she and her team are developing cognitive training programs using technical and communication media (computer, Internet etc.), to improve cognitive abilities in the healthy elderly and those with a mild cognitive impairment to support the Activities of Daily Living (ADL).

Since the project is still very young and the main focus does not address in principal MMI-related problems in terms of IT devices, she could only give some recommendations based on both, her experience in working with elderly (suffering from Alzheimer, Parkinson, visual impairments, restricted capacity to hear, all in different state) and the observation she made during this work.

Dr. Kiss approaches the problems more from the technical side. The goal of his research group is to enhance the quality of life of elderly people by improving and developing already available technologies. One field of application refers to mobility research in order to increase driving safety by "reducing mental workload for elderly drivers in the cockpit of a car", another to "externalization of memory" for people with memory deficits through the development of adapted memory systems such as in palm tops.

Both agreed that technologies and systems have to be intelligent, and that they have to be adapted to the persons affected with their respective information processing capacities and mental abilities.

"Keep it simple" could be the main recommendation, based on the information collected during the mission trip to Bad Tölz. The access to the different functions should be as direct as possible. Hidden functionalities will lead either to confusion or will not be used. The less controls on the surface of the device the better. Interaction that has to be learned first will not be accepted. Displays which provide too much information at the same time lead to misunderstandings and this way to input errors. Speaker independent voice input was therefore a major requirement for future IT devices, of course flanked by appropriate optical and acoustical feedback (good contrast, not too coloured, no high frequencies, easily adjustable volume…). A further requirement, mentioned during the interview, was an easily comprehensible recovery function that helps to get back in idle once lost in undesired sub-menus/functions the user is not familiar with.

As a follow-up to this mission trip to the GRP, usability with elderly was planed and some weeks later conducted. A standard cordless phone was subjected to comparative testing with a novel, especially according to Design for All (DfA) requirements designed other cordless phone. This test could confirm many aspects, mentioned by Dr. Kiss and Mrs. Poulaki. It also showed that supplying a device just with bigger keys and a bigger display font cannot be the solution to compensate the cognitive and sensory-motoric weaknesses of aged users of IT products. In a concluding interview all participants named in principle all requirements that were summarized during the meeting with the member of the GRP.

Since it is matter of fact that the use of IT devices by aged persons is still in the beginning of its development, the outcome of the above-described activities has to be considered as essential. Smart use of multi-modal interaction can help to ensure as long as possible a participation in the benefits of modern communication.

Annex B (informative): Applying the design guidelines to a real world scenario

As an example we have taken the scenario used at the Dundee workshop and extracted example screens from this scenario. The goal for the scenario task was the registration and the online-purchase of a loaf of bread from an online shopping site. After each step of the scenario task we have taken screen shots and we will show what multimodal input and output alternatives and design decision could be taken to make the described transactions multimodal.

Open browser and go to the URL www.tesco.com.

Figure B.1: The start-up screen for the online-shopping system

Principle 1 requires the on this screen the presentation of different input and output modalities selectable by the user, principle 2 requires the means to choose between these modalities using any of the modalities present. According to principle 7 these settings should be persistent and modifiable after any complete transactional step.

Press the shopping tab.

Figure B.2: Choosing different shopping options

According to Principle 3, a selection bar for activating/deactivating various modalities should be present at any time.

Complete the New Customers registration and press register. Then complete the New Customers registration...

Figure B.3: Registering for the shopping application

According to Principle 4, the user should be able to modify the size of the output fonts, or, if audio-output is chosen, the volume of the audible output. Also, if the option to choose different input means is available, the user should be able to switch between these input means after completing input to any of the required input fields (Principle 6).

... and press "accept"

Figure B.4: After completion of the registration process

Go to the local grocery store

Figure B.5: Information on how to use the local grocery delivery

According to Principles 9 and 10 this information should be available in alternate media.

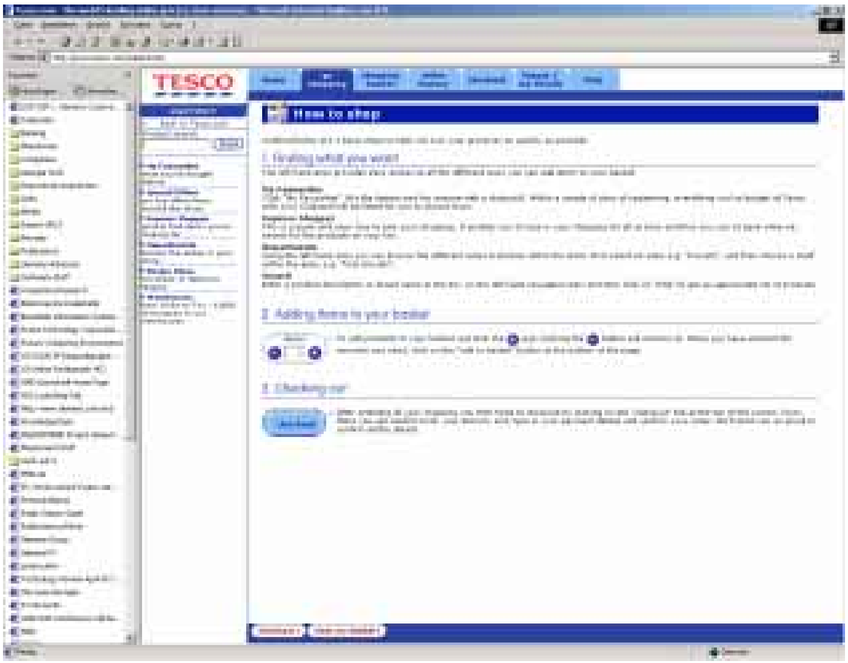

Figure B.6: Help information for the shopping transaction

According to Principle 9 all the information on the preceding screens should be available in different modalities.

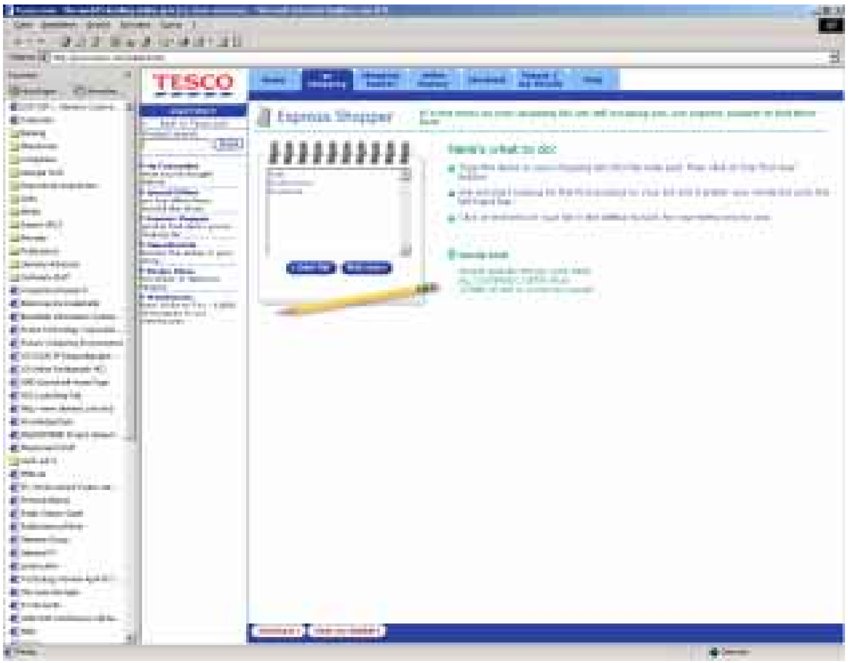

Figure B.7: Description on how to use the virtual notepad

Type bread into the virtual notepad

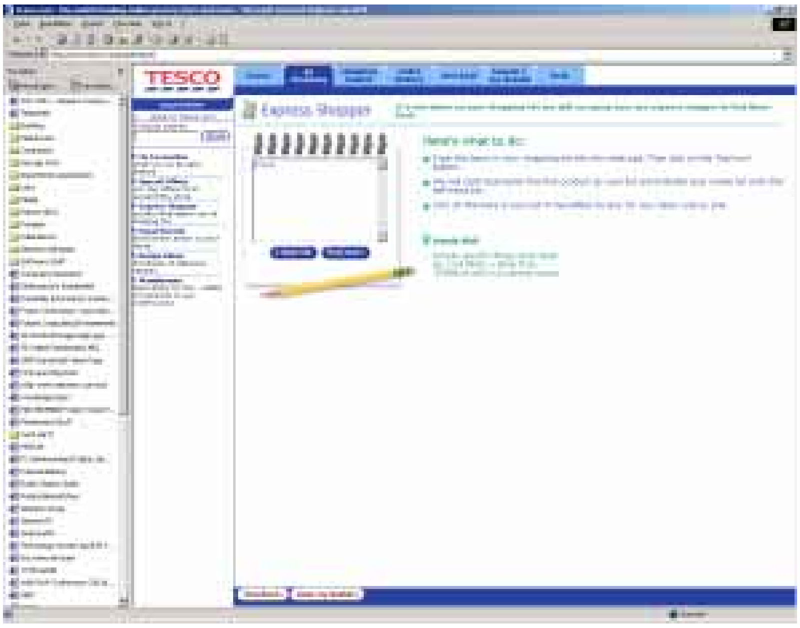

Figure B.8: Choose your grocery items

The list of available "Bread" items

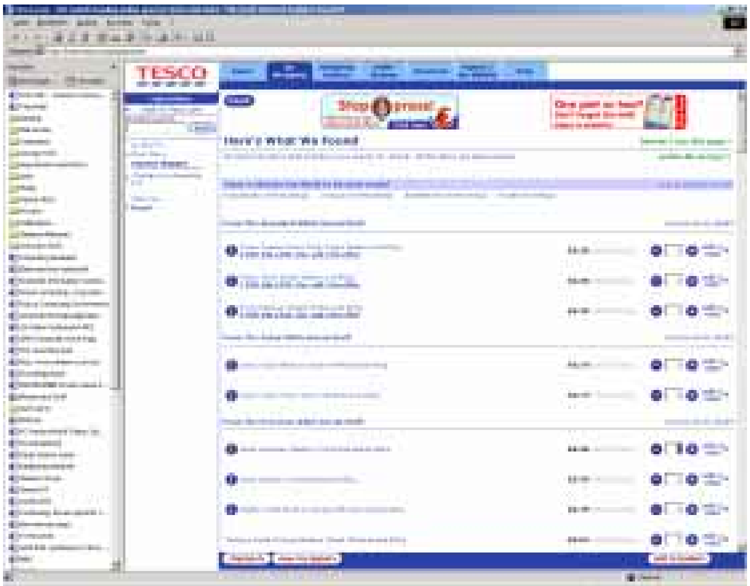

Figure B.9: The result list for a selection

According to Principles 9 and 10 all the information and the selection means should be available in alternate modalities. Also, following Principle 8, prompts should be available to identify the position where user input is required.

Selecting one item

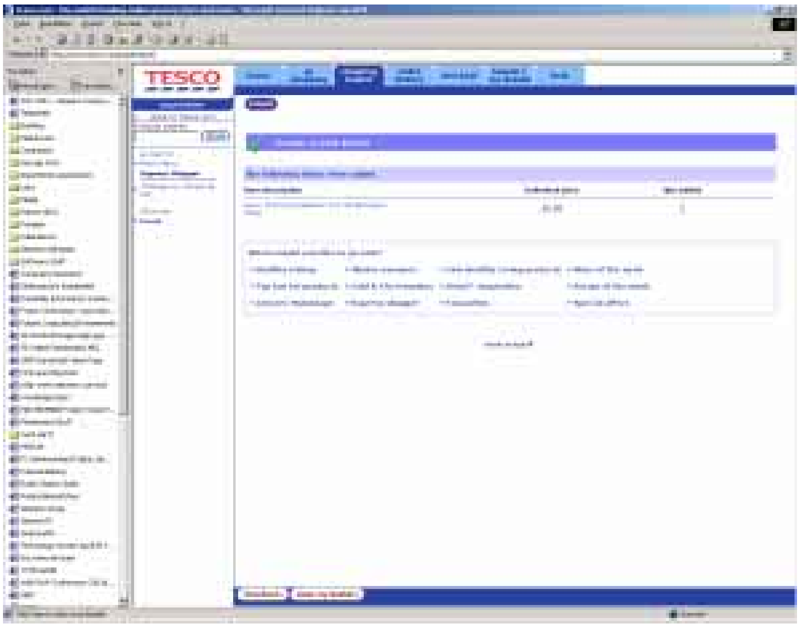

Figure B.10: Final selection

Choose your delivery date

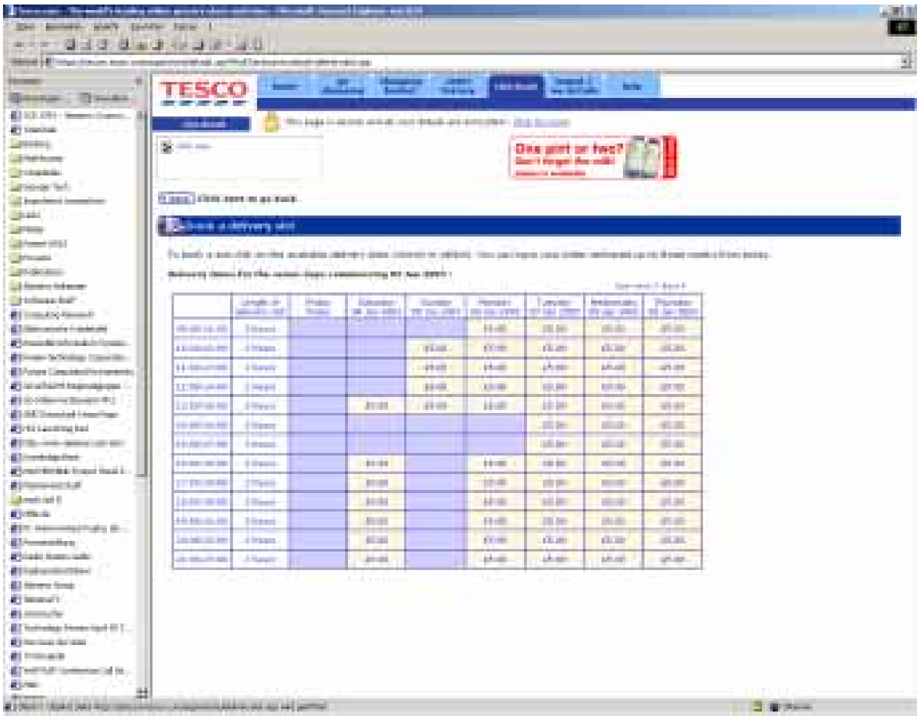

Figure B.11: Selection of the delivery date

According to Principles 9, 10 and 11, adequate means should be provided to efficiently display the information available to users with no or limited eyesight. Therefore an adequate representation of the table displayed must be defined.

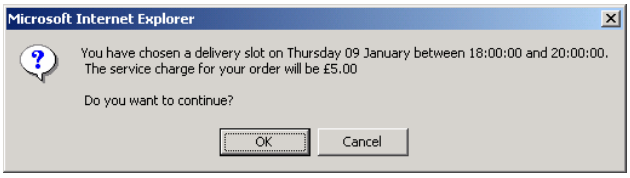

Figure B.12: Feedback on the delivery selection

Confirm delivery

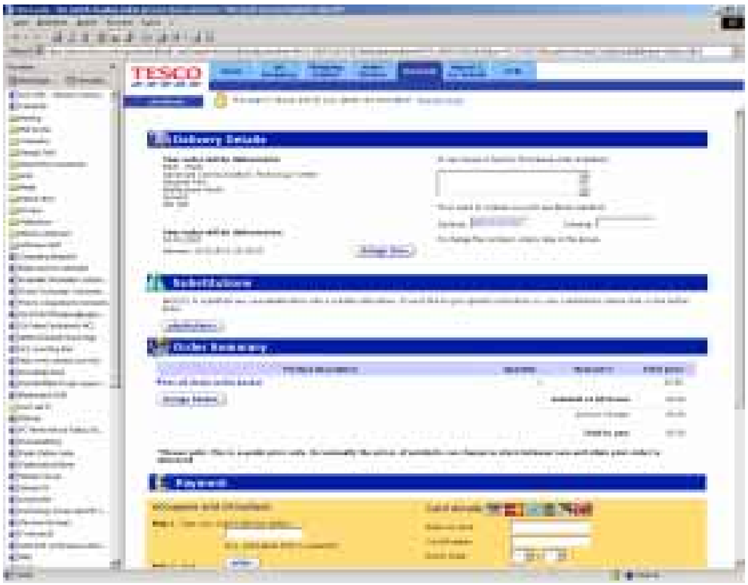

Figure B.13: Confirmation and payment details

According to principle 9 adequate translation means to present the meaning of the colour code used on this screen should be provided for people who are colour blind or rely on audio output.

The available help page

Figure B.14: Help information

Leaving the system

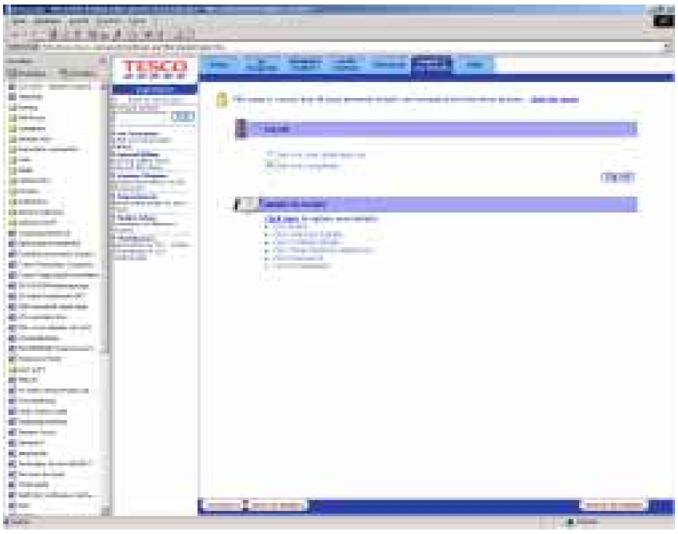

Figure B.15: Exit confirmation

Annex C (informative): Current initiatives (technology roadmap)

This annex catalogues a set of findings and opinions expressed by key researchers in the domains relevant to this study. These findings and opinions were gathered primarily during a mission trip to the USA during April/May 2002.

C.1 World Wide Web Consortium (W3C), Web Accessibility Initiative (WAI)

C.1.1 Judy Brewer - Director, Web Accessibility Initiative, International program Office

Discussed in depth the content accessibility issues as perceived by the W3C. The W3C is involved in the following web accessibility activities:

- ensuring that W3C technologies (HTML, CSS, SMIL, XML, etc.) support accessibility; developing accessibility guidelines for Web content, user agents, authoring tools, and XML applications;
- developing tools for evaluation and repair of Web sites;
- conducting education and outreach on Web accessibility solutions;
- monitoring research and development which may impact the future accessibility of the Web.

An important point here was that in some cases the W3C recommendations are being diluted, most notably in the area of cognitive impairment. This could have profound implications for elderly users of transactional service, and for users who are moving into the community who have been living in institutions for much of their lives.

C.2 Media Access Group at WCBH, National Centre for Accessible Media (NCAM)

C.2.1 Geoff Freed, Project Manager

The Access to Rich Media project was presented. This is concerned with ensuring that these media are annotated or captioned in such a way as to convey their contents to people with visual or auditory impairments. This involves ensuring that the media formats can support additional annotations or captioning, primarily by supporting the SMIL standard. Additional activities consist of the creation of tools such as the Media Access Generator (MAGpie) for easy captioning of rich media, and a web site resource for designers and users of rich media.

A practical example of this type of work that was also presented was the Access to Physics Interactive Video Tutor (PIVoT) project, which intends to make online physics curriculum accessible to deaf and blind students.

C.2.2 Gerry Field, Technology Manager

Solutions were demonstrated for captioning DTV and for providing multiple captioning streams for different captioning requirements. The possibilities for ensuring the accessibility of MPEG-4 video streams were discussed.

C.3 Critical Computing Research Group, MIT Media Lab

Brian Smith

Various aspects of the use of technology in supporting independent living were discussed. The e primary specific research project of interest is the Digital Mirror. In this project, users make multimedia records of their lives and compare that with the telemetry coming from diabetes monitoring equipment to find a correlation between lifestyle and health. This type of project illustrates an interaction between health professionals and patients that can be conducted, to a large degree, remotely.

(http://web.media.mit.edu/~frost/DM.html)

C.4 IBM Research Labs

The IBM Research Lab at Yorktown and Hawthorne NY houses a research group looking at the issues of accessibility that may eventually influence the design of future IBM systems and services. This is distinct from the research teams at Austin TX who advise on the implementers of the current systems.

C.4.1 Shari Trewin

Described her work in developing accessible adaptable user interfaces. This work monitored user performance when typing, made inferences about the user's UI requirements, and adjusted the UI accordingly.

C.4.2 Stéphane Maes - Manager, Mobile Speech Solutions, human language technologies

Gave a thorough overview of the technologies for nomadic device independent access to services, including the capability to undertake the transaction across a variety of devices simultaneously. A reference architecture is being developed and presented to the various standards bodies. The architecture for device independent content handling is based on X-forms and VoiceXML.

C.4.3 Jay Murdock - Industry Solutions Lab, & Teresita Abay-Krueger - DISU Event Manager, Industry Solutions Lab

Demonstrated a secure voice based system for transactional services. The goal is to reduce the high rejection rate of current voiced based secure access systems. This system is based on the recognition of the speech patterns of each user's voice, and on their answers to a set of predefined questions. Uncertainty in the recognition causes additional questions to be asked. For each of these the speech patterns are analysed, and parameters such as the hesitation and uncertainty in answering the questions.

A home shopping system was demonstrated that was developed for a UK supermarket chain. This system was based on a PDA with a barcode scanner. The system was used at home to replenish a shopping list with items before disposing of the packaging, using the barcode scanner or by selecting items from list, or by searching for items from the shop's catalogue. When the PDA is docked on its cradle, the data is transferred from the device to the shop, and updated offers and other user specific selected items are uploaded. The system had been extensively tested with a cross section of the stores more regular customers, but extensive trials with disabled customers had not taken place.

C.4.4 Jennifer Lai - Speech Interface Design, Pervasive Computing

Is responsible for a number of projects involving the use of Speech User Interfaces (SUI). The primary finding of her work has been that there is considerable difficulty in implementing functionally equivalent interfaces that involve making choices from long lists or where the objects being chosen have extensive descriptions. Speech is often used as one representation or interaction modality within a complete interface, but invariably there is little redundancy of information between the modalities.

Another important finding is that when speaking, people are often not rigorous or accurate in their use of grammar. This makes it difficult to use grammar based natural language techniques when analysing spoken interactions. For this reason, current techniques are focussing on the use of statistical methods based on past usage histories of individual users.

One interesting project within this group is the wearable identity device known as the e-pin. This device can be used, alone or in conjunction with another identity system, to provide a secure and easy-to-use means of confirming identifying a user as being in a location.

C.4.5 Rick Kjeldsen - Vision-based user interfaces

Rick demonstrated a number of vision based user interface systems, including a pointing devices that detected head movements (by tracking the user's nose) and tracking of finger pointing. These systems detect movements, and then use IBM's via voice or an alternative modality as the means of making a switch selection.

A more futuristic system, known as the everywhere display, uses projected images as a means of providing a help system, with a camera detecting the interaction between the user and the projected images. One application for this interface that is currently being piloted is assistance to domestic users interacting with home systems and with e-shopping applications.

C.4.6 Vicki Hansen

Described practical aspects of providing accessibility to older people and people with disabilities. When using web-based transaction services, one method that can be used is to change the behaviour of the web browser or client application. On a windows platform this can be achieved by overriding the registry settings for the browser and thereby changing presentational aspects such as size and style of fonts being used &c. It is conceivable that client application behaviour could be changed with methods such as this, based on settings in a user profile that is invoked at login.

An example of Vicki's work is the Universal Accessibility Gateway (UAG).

C.4.7 Medical Automation Research Centre

Robin Felder and Majd Alwan

The Medical Automation Research Centre (MARC) specializes on rehabilitation robotics and smart home technology to assist in medical care, and the social care needs of people living at home. The primary work of interest focuses on rehabilitation robotics and on smart home technology for monitoring health signs and ADL The rehabilitation robotics work involves work on computer vision and on mobile robots for home assistance. The smart homework involves the use of sensors to track movements, activities and health vital signs. (Including aspects such as heart rate monitoring using highly sensitive pressure air pressure sensors attached to air filled "balloons"). The emphasis of this work is prevention of health problems through cheap and simple but effective solutions for installation in homes of people on lower incomes. This group has a working inventory management system based on blue-tooth.

The primary aspect of this work that is of interest to STF 204 is the remote interact between the client/patient and the health services. Work is being done to simplify the interfaces for the users so that they can readily understand the implications of the telemetry being gathered about them, and to interact with the health services at a distance.

C.5 Georgia Tech

Jim Rowan (GATech)

Conducts research to support elderly people living independently. This has focussed on a lifestyle monitoring application to enable care providers (adult children, wardens etc) to monitor Activities of Daily Living (ADL) within the home. Lifestyle patterns are provided visually, but interpretation is not performed automatically. It is the responsibility of the person reading the visuals to look for anomalies in the patterns.

Another area of research is to provide a social interaction mechanism for children and grandparents. This is a flexible multimedia system where media can be annotated and exchanged as an asynchronous conversation. The child interface is facilitated by "Dude" who interacts with the child and encourages them to construct the message to be sent to the grandparent.

These systems are set within the context of the smart home at GATech. This home has a number of systems for detecting ADL. This can be done by location (in a room or going through a door) or by setting sensors in places such as a fridge or mailbox. If activities are being monitored, then a simple interface can simply show if the mail has been collected or the bath taken.

Within the home there is a comprehensive set of video cameras. These are used to detect the presence of people, filtering one person from another and filtering people from other objects or animals. The system does all the processing in the house and therefore all signals leaving the house are effectively telemetry.

The house is equipped with pressure mats on the floor, sensors detecting motion through a door, and sensors detecting the direction of motion through the door. The cabling problem is partially solved by running concealed lighting behind a "gutter" round the tops of the walls in each room. The cabling is placed within these gutters.

A general point about the smart home examples is that they are still at the stage of proving the technology, and are looking at the data that can be gathered. The user interface aspects have not been thoroughly addressed and tested with large groups of users. The issues currently being addressed focus on the acceptance of the sensors and the nature of the data that can be collected. Because of this, there has been little attention paid to the use of these systems by people with cognitive, sensory or physical impairments. There is, however, and expectation that the interaction with the home system will be multimodal and appropriate. The issues of interaction between the people living in the smart environments and the support, care and health services has hardly been addressed yet so guidelines and recommendations cannot currently be found. The importance of this work is that it is suggesting the increasing deployment of smart technology in the lives and experiences of people living in the community.

C.6 TRACE Centre, University of Madison

Greg Vanderheiden

This group has a number of useful initiatives covering various aspects of usability of technology by users with disabilities. The evidence of the work there seems to suggest that there is a move a way from interaction with transactional services being confined to Web based delivery. For example, the group were conducting user trials with a dedicated device for eVoting that ran client software rather than a web interface. Whilst this system was to be used in a polling station rather than remotely, the suggestion was to enhance the client software to enable remote collection of the votes cast by users in an accessible location (hospital or day care centre for example) rather than use a generic web based system.

C.7 Microsoft

Laura Ruby; George Allen; Madelyn Bryant McIntire, Masahiko Kaneko, Marlin Blizinsky and Jessie Tenenbaum

The various initiatives within Microsoft to address accessibility were discussed, focussing on the SALT (Speech Application Language Tags). These are a set of XML tags to extend the HTML framework to handle speech. This approach is part of the ASP.net strategy for future dynamic HTML technologies. It depends on Internet Explorer as the client delivery platform and in that sense is not as generic as the VoiceXML initiative. The SDK includes evolving tools for voice recognition and text-to-speech, with the preference for digitally recorded speech samples rather than synthesized speech.

Other aspects of accessibility that were discussed included the requirement to improve accessibility given the increasing legislative requirements to ensure that education and public information and services are usable by all.

Annex D (informative): Bibliography

ETSI SR 001 996 (V1.1.1): "Human Factors (HF); An annotated bibliography of documents dealing with Human Factors and disability".

ETSI TR 101 767 (V1.1.1): "Human Factors (HF); Symbols to identify telecommunications facilities for deaf and hard of hearing people; Development and evaluation".

ETSI TR 101 806 (V1.1.1): "Human Factors (HF); Guidelines for Telecommunication Relay Services for Text Telephones".

ETSI ETR 029 (1991): "Human Factors (HF); Access to telecommunications for people with special needs Recommendations for improving and adapting telecommunication terminals and services for people with impairments".

ETSI ETR 068 (1993): "Human Factors (HF); European standardization situation of telecommunications facilities for people with special needs".

ETSI ETR 165 (1995): "Human Factors (HF); Recommendation for a tactile identifier on machine readable cards for telecommunication terminals".

ETSI ETR 166 (1995): "Human Factors (HF); Evaluation of telephones for people with special needs; An evaluation method".

ETSI ETS 300 381 (1994): "Telephony for hearing impaired people; Inductive coupling of telephone earphones to hearing aids".

ETSI ETS 300 488 (1996): "Terminal Equipment (TE); Telephony for hearing impaired people; Characteristics of telephone sets that provide additional receiving amplification for the benefit of the hearing impaired".

ETSI ETS 300 679 (1996): "Terminal Equipment (TE); Telephony for the hearing impaired; Electrical coupling of telephone sets to hearing aids".

ETSI EN 301 462 (V1.1.1): "Human Factors (HF); Symbols to identify telecommunications facilities for the deaf and hard of hearing people".

ETSI ETR 333 (1998): "Human Factors (HF); Text Telephony; Basic user requirements and recommendations".

ETSI ETR 334 (1996): "Human Factors (HF); The implications of human ageing for the design of telephone terminals".

ETSI ETR 345 (1997): "Human Factors (HF); Characteristics of telephone keypads and keyboards; Requirements of elderly and disabled people".

The ETSI Aurora STQ, http://portal.etsi.org/stq/kta/DSR/dsr.asp, presence verified 10/2/03

3GPP, http://www.3gpp.org/ presence verified 10/2/03

CSS Mobile Profile, http://www.w3.org/TR/css-mobile presence verified 10/2/03

Device Independence Activity, http://www.w3.org/2001/di/ presence verified 10/2/03

Enduring economic exclusion: Disabled people, income and work (2002) Joseph Rowntree Foundation. http://www.jrf.org.uk/knowledge/findings/socialpolicy/pdf/060.pdf

IMS Guidelines for Developing Accessible Learning Applications, http://www.imsproject.org/accessibility/index.cfm presence verified 10/2/03

Linguistic Structures for e-mail and chat, http://www.ling.lu.se/disseminations/pdf/49/bidrag17.pdf presence verified 10/2/03

M. Clarke, A. Lysley, C. Nicolle and D. Pulson, "World Wide AAC: Developing Internet Services for People using AAC", in the proceedings of the 2002 biennial Conference of the International Society for Augmentative and Alternative Communication, Copenhagen, Denmark, 10th - 15th August 2002.

Making Educational Software and Web Sites Accessible, http://ncam.wgbh.org/cdrom/guideline/ presence verified 10/2/03

Mobile SVG Profiles: SVG Tiny and SVG Basic, http://www.w3.org/TR/SVGMobile/ presence verified 10/2/03

Position Statement for Multi-Modal Access, http://www.w3.org/2002/mmi/2002/siemens-26nov01.pdf presence verified 10/2/03

SALT, http://www.microsoft.com/speech/evaluation/speechtags/ presence verified 10/2/03

Sharon Oviatt (1999) Ten Myths of Multimodal Interaction, Communications of the ACM

SVG 1.0 Specification, http://www.w3.org/TR/SVG/ presence verified 10/2/03

Tim Berners-Lee, James Hendler and Ora Lassila, May 17, 2001, The Semantic Web, Scientific American

Universal Accessibility Gateway (UAG) (http://virtual.inesc.pt/wuauc01/procs/pdfs/hanson_final.pdf and http://www.springerlink.com/app/home/content.)

User Agent Accessibility Guidelines, http://www.w3.org/TR/2002/REC-UAAG10-20021217/ presence verified 10/2/03

Voice Browser Activity, http://www.w3.org/Voice/ presence verified 10/2/03

VoiceXML 2.0 Specification, http://www.w3.org/TR/voicexml20/ presence verified 10/2/03

XForms - The Next Generation of Web Forms, http://www.w3.org/MarkUp/Forms/ presence verified 10/2/03

Disability and Social Exclusion in the European Union: Time for Change, Tools for Change. Study funded by the European Commission DG Employment and Social Affairs (Project Number VP/2000/008-qurey).

Delivering Accessibility: Improving disabled people's assess to the knowledge based Society. Commission Staff Working Paper. Commission of the European Communities, Brussels. SEC(2002) 1039.

eEurope 2005: An information society for all, An Action Plan to be presented in view of the Sevilla European Council, 21/22 June 2002
http://europa.eu.int/information_society/eeurope/news_library/documents/eeurope2005/eeurope2005_en.pdf

The employment situation of people with disabilities in the European Union, a study by EIM Business and Policy Research for European Commission Directorate-General for Employment and Social Affairs Unit EMPL/E.4

History

Document history		
V1.1.1	June 2003	Membership Approval Procedure MV 20030801: 2003-06-03 to 2003-08-01
V1.1.1	August 2003	Publication