Final draft ETSI EG 202 048 V1.1.1 (2002-06)

ETSI Guide

Human Factors (HF);
Guidelines on the multimodality
of icons, symbols and pictograms

Reference

DEG/HF-00027

Keywords

icons, interface, multimode, pictograms, symbols, 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

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), and is now submitted for the ETSI standards Membership Approval Procedure.

Introduction

The *e*Europe initiative

On 8 December 1999 the European Commission launched an initiative entitled "eEurope - An Information Society for All", which proposes ambitious targets to bring the benefits of the Information Society within reach of all Europeans. The initiative focuses on ten priority areas, one of them specifically addressing the needs of people with disabilities. The targets set up by the Communication for improving access to the information society for people with disabilities include the following [15]:

1) A review of the relevant legislation and standards programmes dealing with the Information Society, with a view to ensuring their conformity with accessibility principles and accelerating standardization processes.

2) The Commission will prepare a Communication on how public procurement instruments can positively take account of the needs of people with disabilities in the public procurement of information and communication technology products and services.

3) European Institutions and the Member States should endorse the existing Web Accessibility Initiative (WAI) guidelines, making the design and content of all public Web sites accessible to people with disabilities (http://www.w3.org/tr/wai-webcontent). The document expresses the target to adopt the Web Accessibility Initiative (WAI) guidelines for public websites, by the European Commission and by Member States by the end of 2001 [15].

The present document is the product of a STF (Specialist Task Force) funded within the eEurope initiative and specifically addresses the first of the objectives presented before, with the particular focus on the application of multimodal interfaces to the Information Society applications and services by applying the Design for All approach. It also addresses objective 3 above, as multimodal interfaces are also covered within the WAI initiative.

The need for a Design for All approach

Eurostat statistics show that there are over 375 million people in the 15 countries of the European Union [17]. At any point in time, the number of people in the European Union directly affected by some form of disability has been estimated at around 10 % of the total population. Elderly people are an equally significant sector of the population: over 77 million - 20 per cent, or one in five - are 60 years of age or more. And as a consequence of the demographic change which is happening throughout Europe, this proportion is constantly increasing. Therefore providing ICT products and services to people with some kind of difficulty in their relationship with the environment is no longer a niche market.

The Design for All approach in the development of technological products and services also generally make them easier to use for everyone else, besides maximizing the possibility of having a product accessible to older and disabled people. STF 183 reviewed the available guidelines and legislation on Design for All, and presents this information in annexes A and F.

User requirements for multimodal symbols: Human abilities and disabilities

People with disabilities and elderly people do not constitute a homogeneous group and there is a broad range of disabilities and other limitations that have to be considered. Disabilities may be apparent or hidden, severe or mild, singular or multiple, chronic or intermittent. Types of disabilities include mobility/agility, mental/cognitive, hearing, speaking, and visual impairments. STF 183 reviewed the user requirements for multimodal interfaces, with special focus on a broad range of disabilities, and some information is provided in the informative annex B of the present document.

The Design for All approach and the development of multimodal interfaces

Multimodal interfaces can achieve more natural and effective human-computer interaction by integrating a variety of signals by which humans usually convey information. The fact that the world is multimodal means that the recognition of a unimodal symbol may be hampered, even if it represents a real world equivalence but in isolation.

The explicit need to develop a multimodal interface may stem from several requirements. These requirements can, amongst others, be related to the user, to the task demands, to the character of the information, or to environmental constraints. An example of a high task demand is reading an SMS message while driving. Also, environmental constraints may necessitate an additional modality. Examples include the limited visual displays of mobile devices, noisy environments, and the field of tension between public and private information presentation (e.g. a ringbone vs. a vibration to indicate an incoming call).

The effectiveness of the user interface to adapt to these many situations and constraints is based on redundancy, i.e. providing information (that is otherwise not available or in a degraded quality) to an alternative or additional information channel. This redundancy may be beneficial for all users and across all situations. The use of multimodal information is then one of the key solutions in a Design for All approach to ICT design. Many of the needs of people with sensory disabilities, cognitive disabilities, and elderly people can be met within mainstream products if they are able to provide symbols and other output information in a range of different modalities. For example, the needs of blind and dyslexic people can be met if text and graphic information is also provided in synthetic speech output and the needs of hearing impaired people can be met if speech and sound information is also provided in text.

Many of the technical possibilities for multimodal interfaces are existing from long ago, and are the subject of specific guidelines and standards, for instance, systems based on speech recognition, and auditory interfaces incorporated in many devices and services. Currently, the most interesting aspect is the technological developments around tactile display elements, that will soon lead to the possibilities of introducing active displays on a large scale. Nowadays, using vibration in mobile phones and pagers is common practise, just as the use of a vibrating computer mouse. These active tactile displays add the possibility of using the tactile sense as a full channel in the man-machine interface. The potential information transfer capacity of this channel is much larger than the 1-bit message "your phone is ringing". It may be foreseen that the use of the tactile modality will grow fast. Not only as an alternative information channel for people with special needs, but for all possible users.

At this moment, there is a lack of guidelines on the design and application of active tactile displays (with the exception of Braille displays and vibrating alerting devices), and on the interaction with the other sensory modalities (multimodal interaction). Since the pool of best practises is also very limited, the most important source to distil the guidelines from is the neurophysiological and psychophysical literature.

The purpose of the Human Factors guidelines in clause 5 of the present document is to review and provide in a concise way the basic relevant information for designers of multimodal user interfaces, with the added value of providing guidelines that were not previously made in this field.

The reader should note that the status of the work on multimodal symbols and interfaces is in an emerging phase. That means that most of the guidelines are tentative and have not been thoroughly validated, or might not be directly applicable to a particular purpose or device. However, we strongly believe that even in the present, preliminary, form they can be useful; both to the designer and to identify interesting areas of future research.

The Human Factors guidelines for multimodal symbols are provided in clause 5 of the present document. It has two main parts:

- In clause 5.1, we start by presenting unimodal guidelines, ordered to the five human senses.
- In clause 5.2, we integrate the unimodal guidelines, discuss specific multimodal effects, and formulate the multimodal guidelines.

Finally examples of the application of these guidelines in real products and of best practice for the successful application of the guidelines are provided in annex D.

1 Scope

The present document presents guidelines for the design and use of multimodal symbols using a Design for All approach. It also provides a study of the needs and requirements for the use of multimodal symbols in user interfaces, with special emphasis on the requirements of people with disabilities and elderly people.

The present document provides guidelines, good practice and case studies for the successful design and application of multimodal symbols using the Design for All approach. The present document will support the standardization process with respect to the use of multimodal symbols in modern user interfaces.

Icons, symbols and pictograms are widely used components of user interfaces in ICT applications and services, e.g. for navigation, status indication and function invocation. Examples of such applications and services include information retrieval (e.g. Web sites), messaging (e.g. email and SMS), public services (e.g. public telephones and ATMs) and real time communication services (e.g. fixed and mobile telephony).

The use of visual-only symbols in such applications and services creates temporary or permanent problems for all users. User groups most affected are blind and partially sighted people and users of mobile devices with limited visual display capabilities. All users can potentially benefit from the current and future possibilities of multimodal user interfaces. These interfaces combine communication channels, for example sound, graphics, video, speech, force and vibration.

The present document does not deal with unimodal symbols, i.e. only visual symbols or only auditory "earcons", but with symbols that use at least two communication channels.

The applications and services make use of many different networks for communication, but this aspect is outside the scope of the project.

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.

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

[2] ETSI ETS 300 375: "Human Factors (HF); Pictograms for point-to-point videotelephony".

[3] ETSI ETR 070: "Human Factors (HF); The Multiple Index Approach (MIA) for the evaluation of pictograms".

[4] ETSI ETR 113: "Human Factors (HF); Results of an evaluation study of pictograms for point-to-point videotelephony".

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

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

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

[8] ETSI ETS 300 767: "Human Factors (HF); Telephone Prepayment Cards; Tactile Identifier".

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

[10] ETSI EG 201 379: "Human Factors (HF); Framework for the development, evaluation and selection of graphical symbols".

[11] ETSI ES 201 381: "Human Factors (HF); Telecommunications keypads and keyboards; Tactile identifiers".

[12] ITU-T Recommendation E.121 (1995): "Pictograms, symbols and icons to assist users of the telephone service".

[13] ITU-T Recommendation F.910 (1995): "Procedures for designing, evaluating and selecting symbols, pictograms and icons".

[14] ETSI EG 201 013: "Human Factors (HF); Definitions, abbreviations and symbols".

[15] eEurope, An Information Society For All: "Progress report: for the Special European Council on Employment, Economic reforms and Social Cohesion - Towards a Europe based on Innovation and Knowledge. Commission of the European Communities. Brussels, 8/3/2000".

[16] ISO/IEC GUIDE 71 (2001): "Guidelines for standards developers to address the needs of older persons and persons with disabilities".

[17] Towards a Barrier Free Europe for People with Disabilities: "communication from the Commission to the Council, the European Parliament, the economic and social committee and the committee of the regions, Commission of the European Communities, Brussels, 12.05.2000".

[18] ICTSB Project Team, Design for All. Final Report, May 2000.

[19] ISO/IEC 11581-1 (2000): "Information technology - User system interfaces and symbols - Icons symbols and functions - Part 1: Icons - General".

[20] ISO/IEC 11581-2 (2000): "Information technology - User system interfaces and symbols - Icon symbols and functions - Part 2: Object icons".

[21] ISO/IEC 11581-3 (2000): "Information technology - User system interfaces and symbols - Icon symbols and functions - Part 3: Pointer icons".

[22] Void.

[23] Void.

[24] ISO/IEC 11581-6 (2000): "Information technology - User system interfaces and symbols - Icon symbols and functions - Part 6: Action icons".

[25] ISO/IEC 80416-1 (2001): "Basic principles for graphical symbols for use on equipment - Part 1: Creation of symbol originals".

[26] ISO/IEC 80416-2: "Basic principles for graphical symbols for use on equipment - Part 2: Form and use of arrows" (to be published)".

[27] ISO 13407 (1999): "Human-centred design processes for interactive systems".

[28] Void.

[29] ETSI EG 201 472: "Human Factors (HF); Usability evaluation for the design of telecommunication systems, services and terminals".

[30] ETSI ETR 116: "Human Factors (HF); Human factors guidelines for ISDN Terminal equipment design".

[31] Akamatsu, M & Sato, S. (1992): "Mouse type interface device with tactile and force display - multimodal integrative mouse", *Proceedings of the second international conference on artificial reality and tele-existence*, pp. 177-182.

[32] Akamatsu, M. & Sato, S. (1992): "The multimodal integrative Mouse - a mouse with tactile display", Proceedings of ACM CHI '92 Conference on human factors in computing systems".

[33] Ananova (2001): "Ananova - http://www.ananova.com/", Ananova Ltd., Available: HYPERLINK "http://www.ananova.com/"/(25 September 2001)".

[34] Baecker, R., Small, I., & Mander, R. (1991): "Bringing Icons to Life", *Proceedings of ACM CHI'91 Conference on Human Factors in Computing System*, pp. 1-6.

[35] Barfield, W., Rosenberg, C. and Levasseur, G. (1991): "The use of icons, earcons, and commands in the design of an online hierarchical menu, *IEEE Transactions on Professional Communication, 34(2)*, pp. 101-108.

[36] Beauregard, G.L. & Srinivasan, M.A. (1997): "Sensorimotor interactions in the haptic perception of virtual objects" (RLE Technical Report No. 607), Cambridge, MA: Massachusetts Institute of Electronics".

[37] Beauregard, G.L. Tan, M.A. & Durlach, N.I. (1995): "The manual resolution of viscosity and mass", *Proceedings of the ASME Dynamic Systems and Control Division*, vol. 57(2), pp 657-662.

[38] Behar, I. & Bevan, W. (1961): "The perceived duration of auditiory and visual intervals: Cross-modal comparison and interaction", *American Journal of Psychology, 74*, pp 17-26.

[39] Berglund, B., Preis, A. and Rankin, K. (1990): "Relationship between loudness and annoyance for ten community sounds", *Environment International, 16*, pp 523-531.

[40] Blattner, M. M., Sumikawa, D.A. and Greenberg, R.M. (1989): "Earcons and icons: their structure and common design principles", *Human Computer Interaction, 4*, pp 11-44.

[41] Bliss, J.C., Crane, H.D., Link, S.W. & Townsend, J.T. (1966): "Tactile perception of sequentially presented spatial patterns, *Perception & Psychophysics, 1*, pp 125-130.

[42] Bolanowski, S.J., Gescheider, G.A. & Verrillo, R.T. (1994): "Hairy skin: Psychophysical channels and their physiological substrates, *Somatosensory Motor Research, 11*, pp 279-290.

[43] Bolanowski, S.J., Gescheider, G.A., Verrillo, R.T. & Checkosky, C.M. (1988): "Four channels mediate the mechanical aspects of touch", *Journal of Acoustical Society of America*, 84(5), pp 1680-1694.

[44] Bonsor, K.: "How Internet Odors Will Work", Marshall Brain's HowStuffWorks, http://www.howstuffworks.com/internet-odor.htm?printable=1, (2001, 25 September)".

[45] Brewster, S. (1994): "Providing a structured method for integrating non-speec haudio into human-computer interfaces", Phd Thesis, Computer Science Department, York, University of York".

[46] Brewster, S. (1998): "Using nonspeech sounds to provide navigation cues", *ACM Transactions on Computer-Human Interaction*, 5, pp 224-259.

[47] Brewster, S. A., Wright, P. C., and Edwards, A. D. N. (1994): "The Design and Evaluation of an Auditory-Enhanced ScrollBar", Proceedings of ACM CHI'94 Conference on Human Factors in Computing Systems (Vol. 1, pp. 173-179)".

[48] Brewster, S. A., Wright, P. C., Dix, A. J., and Edwards, A. D. N. (1995): "The sonic enhancement of graphical buttons".

[49] Brewster, S., Wright, P. C. and Edwards, A.D.N.: "Parallel earcons: reducing the length of audio messages", *International Journal of Human-Computer Studies*, 43(2), pp 153 - 175".

[50] Brewster, S., Wright, P.C. and Edwards, A.D.N. (1992): "A detailed investigation into the effectiveness of earcons", Proceedings of the First International Conference on Auditory Display (ICAD '92), Santa Fe, NM, Addison-Wesley.

[51] Brewster, S.A., Wright, P.C. & Edwards, A.D.N. (1994): "The desing and evaluation of an auditory- enhanced scrollbar", Proceedings of ACM CHI '94 Conference on human factors in computing systems.

[52] Brodie, E.E. & Ross, H.E. (1984): "Sensorimotor mechanisms in weight discrimination", *Perception & Psychophysics*, 36(5), pp 477-481".

[53] Brodie, E.E. (1988): "Sex and hand-preference effects in the simultaneous and consecutive discrimination of lifted weight", *Perception & Psychophysics*, 43(4), 326-330".

[54] Buxton, W., Gaver, W.W. and Bly, S. (1991): "The use of non-speech audio at the interface", Proceedings of CHI '91, New York: ACM Press".

[55] Cain, W.S. and Stevens, J.C. (1971): "Effort in sustained and phasic handgrip contractions", *American Journal of Psychology, 84*, pp 52-65.

[56] Cain, W.S. (1973): "Nature of perceived effort and fatigue: roles of strength and blood flow in muscle contractions", *Journal of Motor Behavior*, 5, pp 33-47".

[57] Capraro, A.J., Verrillo, R.T. & Zwislocki, J.J. (1979): "Psychophysical evidence for a triplex system of cutaneous mechanoreception", *Sensory Processes, 3*, pp 340-352.

[58] Carlson, N.K., Drury, C.G. & Webber, J.A. (1977): "Discriminability of large weights", *Ergonomics, 20(1)*, pp 87-90.

[59] Cholewiak, R.W. (1979): "Spatial factors in the perceived intensity of vibrotactile patterns", *Sensory Processes, 3*, pp 141-156.

[60] Cholewiak, R. and Collins, A. (1991): "Sensory and physiological bases of touch".

[61] Cohen, B. & Kirman, J.H. (1986): "Vibrotactile frequency discrimination at short durations", *Journal of General Psychology, 113(2)*, pp 179-186.

[62] Colwell, C., Petrie, H., Kornbrot, D., Hardwick, A. and Furner, S. (1998): "Haptic virtual reality for blind computer users", Proceedings of ASSETS '98: The Third International ACM Conference on Assistive Technologies, New York: ACM Press. ISBN 1-58113-020-1".

[63] Conolly, K. & Jones, B. (1970): "A developmental study of afferent-reafferent integration", *British Journal of Psychology*, 61, pp 259 - 266.

[64] Cook, M.J., Angus, C., Campbell, C. & Cranner, C., (1999): "Speech interfaces in real-time control systems", *People in control. London: Institution of electrical engineeers*, pp. 428-433.

[65] Cook, M.J., Cranner, C., Finan, R., Sapeluk, A. and Milton, C.A. (1997): "Memory load and task interference: hidden usability issues in speech interfaces", in: D. Harris (ed.) Engineering psychology and cognitive ergonomics, part 1. Aldershot, UK: Ashgate publishing, pp. 141-150.

[66] Cook, S.: "Next on the "virtual" horizon: scents", *The Christian Science Monitor*. Available: http://www.csmonitor.com/durable/2001/03/12/fp14s1-csm.shtml, 25 September 2001".

[67] Cooper, D.F., Grimby, G., Jones, D.A. & Edwards, R.H.T. (1979): "Perception of effort in isometric and dynamic muscular contraction", *European Journal of Applied Physiology, 41*, pp 173-180.

[68] Craig, J.C. (1968): "Vibrotactile spatial summation", *Perception & Psychophysics, 4(6)*, pp 351-354.

[69] Craig, J.C. (1972): "Difference threshold for intensity of tactile stimuli", *Perception & Psychophysics, 11(2)*, pp 150-152.

[70] Craig, J.C. (1974): "Vibrotactile difference thresholds for intensity and the effect of a masking stimulus", *Perception & Psychophysics, 15(1)*, pp 123-127.

[71] Craig, J.C. (1982): "Vibrotactile masking: A comparison of energy and pattern maskers", *Perception & Psychophysics, 31(6)*, pp 523-529.

[72] Craig, J.C. (1982b): "Temporal integration of vibrotactile patterns", *Perception & Psychophysics, 32(3)*, pp 219-229.

[73] Craig, J.C. (1989): "Interference in localizing tactile stimuli", *Perception & Psychophysics*, 45(5), pp 343-355".

[74] Craig, J.C. and Baihua, X. (1990): "Temporal order and tactile patterns", *Perception & Psychophysics, 47(1),* pp 22-34.

[75] Craig, J.C. and Evans, P.M. (1987): "Vibrotactile masking and the persistence of tactual features", *Perception & Psychophysics, 42(4),* pp 309-317.

[76] Craig, J.C. and Sherrick, C.E. (1969): "The role of skin coupling in the determination of vibrotactile spatial summation", *Perception & Psychophysics*, *6(2),* pp 97-101.

[77] Darwood, J.J., Repperger, D.W. and Goodyear, C.D. (1991): "Mass discrimination under G_z acceleration", *Aviation, Space, and Environmental Medicine, 62(4),* pp. 319-324.

[78] Davidson, R. & Mather, J.H. (1966): "Cross-modal judgments of length", *American Journal of Psychology*, 79, pp 409 - 418.

[79] Ehrensing, R.H. & Lhamon, W.T. (1966): "Comparison of tactile and auditory time judgments", Perceptual & Motor Skills, 23, pp 929 - 930.

[80] Digimask (2001): "Digimask" [Web document]. Digimask. Available: "http://www.digimask.com/" (25 September 2001).

[81] Void.

[82] Dormann, C. (1994): "The Design of Animated Signs as Help", Proceedings of ACM CHI'94 Conference on Human Factors in Computing Systems, (Vol. 2, pp. 137-138).

[83] Edworthy, J., Loxley, D. and Dennis, I. (1991): "Improving auditory warning design: relationships between warning sound parameters and perceived urgency", *Human Factors*, 33(2), pp 205 - 231".

[84] Ehrensing, R.H. & Lhamon, W.T. (1966): "Comparison of tactile and auditory time judgments", *Perceptual & Motor Skills*, 23, pp. 929 - 930.

[85] Eisler, H. (1962): "Subjective scale of force for a large muscle group", *Journal of Experimental Psychology*, 64, pp 253-257".

[86] Eisler, H. (1965): "The ceiling of psychophysical power functions", *American Journal of Psychology*, 78, pp 506-509".

[87] Evans, P.M. (1987): "Vibrotactile masking: temporal integration, persistence, and strengths of representations", *Perception & Psychophysics, 42(6),* pp 515-525.

[88] Forren, M.G. & Mitchell, C.M. (1986): "Multimodal interfaces in supervisory control", *Proceedings of the human factors society 30th annual meetin*, pp 317-321.

[89] Freides, D. (1974): "Human information processing and sensory modality: cross-modal functions, information complexity, memory, and deficit", Psychological Bulletin, 81, 5, pp 284-310".

[90] Furner, S., Hardwick, A., Colwell, C., Bruns, T., Petrie, H. and Kornbrot, D. (1999): "Computer haptics: putting physical sensation into the "look and feel" of human computer interaction at the desktop", *British Telecommunications Technical Journal: Special issue on HCI*".

[91] Gaver, W. W. (1986): "Auditory icons: using sound in computer interfaces", Human Computer Interation, 2, pp 167-177".

[92] Gaver, W. W. (1991): "Technology affordances", Proceedings of CHI '91. New York: ACM Press.

[93] Gaver, W. W. (1993): "What in the world do we hear? An ecological approach to auditory event perception", *Ecological Psychology 5*, pp 1 - 29.

[94] Gaver, W. W. (1997): "Auditory Interfaces". In M. Helander & T. K. Landauer & P. V. Prabhu (Eds.), *Handbook of Human-Computer Interaction* (Second, completely revised ed., pp. 1003-1041), Amsterdam: North-Holland Elsevier Science Publishers.

[95] Geldard, F.A. (1970): "Vision, audition, and beyond". In: W.D. Neff (ed.), *Contributions to sensory psychology*, pp. 1 - 17, New York: Academic press.

[96] Geldard, F.A. & Sherrick, C.E. (1972): "The cutaneous rabbit: a perceptual illusion", *Science*, vol. 178, 13 oct. 1972, p. 178.

[97] Gentner, D. (1983): "Structure mapping: a theoretical framework for analogy", *Cognitive Science, 7,* pp 155-170.

[98] Gescheider, G.A. (1974): "Temporal relations in cutaneous stimulation", In: *Cutaneous communication systems and devices*, edited by F.A. Geldard. Austin, Tex.: The Psychonomic Society.

[99] Gescheider, G.A. & Joelson, J.M. (1983): "Vibrotactile temporal summation for threshold and suprathreshold levels of stimulation", *Perception & Psychophysics, 33(2),* pp 156-162.

[100] Gescheider, G.A. & Verillo, R.T. (1979): "Vibrotactiel frequency characteristics as determined by adaption and masking procedures". Ib Kenshalo, D.R. (eds), *Sensory functions of the skin of humans, proceedings of the second international symposium on skin senses*, pp 183-205, New York: Plenum Press.

[101] Gibson, J.J. (1979): "The ecological approach to visual perception", New York: Houghton Mifflin.

[102] Goble, A.K. & Hollins, M. (1994): "Vibrotactile adaptation enhances frequency discrimination", *Journal of Acoustical Society of America, 96(2),* pp 771-779.

[103] Goble, A.K. & Hollins, M. (1993): "Vibrotactile adaptation enhances amplitude discrimination", *Journal of Acoustical Society of America, 93(1),* pp 418-424.

[104] Goff, G.D. (1967): "Differential discrimination of frequency of cutaneous mechanical vibration", *Journal of Experimental Psychology, 74(2)*, pp 294-299.

[105] Goldstone, S. & Goldfarb, J. (1963): "Judgment of filled and unfilled durations: Intersensory effects", *Perceptual & Motor Skills*, *17*, pp 763 - 774.

[106] Goldstone, S. & Lhamon, W.T. (1972): "Auditory-visual differences in human judgment", *Perceptual and Motor skills, 34*, pp 623 - 633.

[107] Goodfellow, L.D. (1934): "An empirical comparison of audition, vision, and touch in the discrimination of short intervals of time", *American Journal of psychology, 46,* pp 243-258.

[108] Green, B.G. (1976): "Vibrotactile temporal summation: Effect of frequency", *Sensory Processes, 1*, pp 138-149.

[109] Green, B.G. (1977): "The effect of skin temperature on vibrotactile sensitivity", *Perception & Psychophysics, 21(3)*, pp 243-248.

[110] Green, B.G. & Craig, J.C. (1974): "The roles of vibration amplitude and static force in vibrotactile spatial summation", *Perception & Psychophysics*, 16(3), pp 503-507.

[111] Hahn, J.F. (1966): "Vibrotactile adaptation and recovery measured by two methods", *Journal of Experimental Psychology, 71(5),* pp 655-658.

[112] Hahn, J.F. (1968): "Low-frequency vibrotactile adaptation", *Journal of Experimental Psychology, 78(4)*, pp 655-659.

[113] Handel, S. (1989): "Listening: an introduction to the perception of auditory events", Cambridge, MA: MIT Press.

[114] Handel, S. & L. Buffardi (1969): "Using several modalities to perceive one temporal pattern", *Quarterly Journal of Experimental Psychology. 21,* pp 256-266.

[115] Hankinson, J. C. K., and Edwards, A. D. N. (1999): "Designing Earcons with Musical Grammars", *ACM SIGCAPH Newsletter, 65*(September), pp 16-20.

[116] Haptek. (2001): "Haptek - www.haptek.com", Haptek Inc. Available: http://www.haptek.com/ (2001, 25 September).

[117] Hawkes, G.R., Deardorff, P.A. & Ray, W.S. (1977): "Response delay effects with cross-modality duration judgments", *The journal of auditory research, 17,* pp 55-57.

[118] Hay, J.C., Pick, H.L. Jr. & Ikeda, K. (1965): "Visual capture produced by prism spectacles", *Psychonomic Science, 2,* pp 215-216.

[119] Heller, M.A. (1991): "Introduction", in M.A. Heller and W. Schiff (Eds.), *The psychology of touch*, Hillsdale, NJ: Lawrence Erlbaum.

[120] Hellström, Å. (2000): "Sensation weighting in comparison and discrimination of heaviness", *Journal of Experimental Psychology: Human Perception and Performance, 26(1),* pp 6-17.

[121] Hirsh, I.J. & Sherrick, C.E. (1961): "Perceived order in different sense modalities", *Journal of Experimental Psychology,* 62(5), pp 423-432.

[122] Hollins, M., Goble, A.K., Whitsel, B.L. & Tommerdahl, M. (1990): "Time course and action spectrum of vibrotactile adaptation", *Somatosensory and Motor Research, 7(2),* pp 205-221.

[123] Jansson, G., Petrie, H., Colwell, C., Kornbrot, D., Fänger, J., König, H., Billberger, K., Hardwick, A. and Furner, S. (1999): "Haptic virtual environments for blind people: exploratory experiments with two devices", *International Journal of Virtual Reality, 4*(1), pp 10 - 20.

[124] Johansson, S.R. (1978): "Tactile sensibility in the human hand: receptive field characteristics of mechanoreceptive units in the glabrous skin area", *Journal of Physiology, 281*, pp 101-123.

[125] Johansson, S.R., Landstrom, U. & Lundstrom, R. (1982): "Responses of mechanoreceptive afferent units in the glabrous skin of the human hand to sinusoidal skin displacements", *Brain Research, 244,* pp 17-25.

[126] Johnson, K.O. & Phillips, J.R. (1981): "Tactile spatial resolution. I. Two point discrimination, gap detection, grating resolution, and letter recognition", *Journal of Neurophysiology, 46(6),* pp 1177-1191.

[127] Jones, L., Hunter, I. & Lafontaine, S. (1997): "Viscosity discrimination: a comparison of an adaptive two-alternative forced-choice and an adjustment procedure", *Perception, 26,* pp 1571-1578.

[128] Jones, L.A. & Hunter, I.W. (1982): "The relation of muscle force and EMG to perceived force in human finger flexors", *European Journal of Applied Physiology, 50,* pp 125-131.

[129] Jones, L.A. & Hunter, I.W. (1990): "A perceptual analysis of stiffness. *Experimental Brain Research, 79,* pp 150-156.

[130] Jones, L.A. & Hunter, I.W. (1993): "A perceptual analysis of viscosity", *Experimental Brain Research, 94,* pp 343-351.

[131] Jones, L.A. (1989): "Matching forces: constant errors and differential thresholds", *Perception, 18,* pp 681-687.

[132] Kahn, J. (2001): "Aroma Therapy In The Military, It's Known as "Nonlethal Weapons Development"" (May 22, 2001), [Web document]. SFGate. Available: http://www.sfgate.com/cgi-bin/article.cgi?file=/gate/archive/2001/05/22/smell.DTL&type=printable (2001, 25 September).

[133] Kandel, E.R., Schwartz, J.H. & Jessel, T.M. (1991): "Principles of Neural Science", (New York: Elsevier Science Publishing Co.).

[134] Kaye, J. N. (2001): "Symbolic Olfactory Display", *Unpublished Master of Science*, Massachusetts Institute of Technology, Cambridge, Boston, MA.

[135] Kirman, J.H. (1973): "Tactile communication of speech: A review and analysis", *Psychological bulletin, 80,* pp 54 - 74.

[136] Kirman, J.H. (1974): "Tactile apparent movement: The effect of interstimulus onset interval and stimulus duration", *Perception and Psychophysics, 15(1)*, pp 1-6.

[137] Kolers, P.A. & Brewster, J.M. (1985): "Rythems and Responses", *Journal of Experimental Psychology: Human Perception and Performance, 11*, pp 150-167.

[138] Lakatos, S. & Shepard, R.N. (1997): "Constraints common to apparent motion in visual, tactile, and auditory space", *Journal of Experimental Psychology, 23(4)*, pp 1050-1060.

[139] Lechelt, E.C. (1975): "Temporal numerosity discrimination: Inter–modal comparsions revisited", *British Journal of Psychology, 66*, pp 101-108.

[140] Lederman, S.J. (1979): "Auditory texture perception", *Perception, 8*, pp 93-103.

[141] Lefcowitz, E. (2001): "Retrofuture - Smell-o-vision!", [Web document], available: http://retrofuture.com/smell-o-vision.html (2001, 25 September).

[142] Ludi, S. (2000): "Animated Icons: Re-Inventing Visual Cues for the Visually Impaired Computer User", *Proceedings of the 2000 International Conference on Intelligent User Interfaces*, pp. 145-146.

[143] MacLean, K. E. and J. B. Roderick (1999): "Smart tangible displays in the everyday world: a haptic door knob", *Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM '99)*, Atlanta, GA.

[144] Makoes, J.C., Gescheider, G.A. & Bolanowski, S.J. (1996): "The effects of static skin indentation on vibrotactile threshold", *Journal of Acoustical Society of America, 99(5)*, pp 3149-3153.

[145] Millar, S. (1972): "The development of visual and kinesthetic judgments of distance", *British Journal of Psychology, 63*, pp 271 - 282.

[146] Miller, T. and Zeleznik, R. (1998): "An insidious haptic invasion: adding force feedback to the X desktop", *Proceedings of UIST '98*, New York: ACM Press.

[147] Miller, T. and Zeleznik, R. (1999): "The design of 3D haptic widgets", *Symposium on Interactive 3D Graphics*, New York: ACM Press.

[148] Mynatt, E. D. and Edwards, W. K. (1992): "Mapping GUIs to Auditory Interfaces", *Proceedings of the ACM Symposium on User Interface Software and Technology*, pp. 61-7.

[149] O'Mara, S., Rowe, M.J. & Tarvin R.P.C. (1988): "Neural mechanisms in vibrotactle adaptation". *Journal of Neurophysiology, 59*, pp 607-622.

[150] Ottensmeyer, M. (1997): "Hot and cold running VR: adding thermal stimuil to the haptic experience", *Proceedings of the Second Phantom Users Group Workshop*, Dedham, MA.

[151] Pang, X.D., Tan, H.Z. and Durlach, N.I. (1991): "Manual discrimination of force using active finger motion", *Perception & Psychophysics, 49*, pp 531-540.

[152] Patterson, R.D. (1982): "Guidelines for Auditory Warning Systems on Civil Aircraft", London: Civil Aviation Authority.

[153] Penn, P., Petrie, H., Colwell, C., Kornbrot, D., Furner, S. and Hardwick, A. (2001): "The perception of virtual textures and objects", in R. Murray-Smith and S. Brewseter (Eds.), *Lecture Notes in Computer Science: Proceedings of First Workshop on Haptic Human Computer Interaction*, Heidelberg: Springer Verlag.

[154] Petrosino, L. and Fucci, D. (1989): "Temporal resolution of the aging tactile sensory system", *Perceptual and Motor Skills, 68*, pp 288-290.

[155] Pick, H.L.Jr., Warren, D.H. & Hay, J.C. (1969): "Sensory conflicts in judgments of spatial direction", *Perception & Psychophysics, 6*, pp 203-205.

[156] Ramstein, C., Martial, O., Dufresne, A., Carignan, M., Chassé, P. and Mabilleau, P. (1996): "Touching and hearing GUIs: Design issues for the PC-access system. In Proceedings of ACM Assets '96 (Vancouver, Canada)", New York: ACM Press.

[157] Roland, P.E. & Ladegaard-Pedersen, H. (1977): "A quantitative analysis of sensations of tension and of kinaesthesia in man", *Brain, 100,* pp 671-692.

[158] Ross, H.E. & Reschke, M.F. (1982): "Mass estimation and discrimination during brief periods of zero gravity", *Perception & Psychophysics*, 31, pp 429-436.

[159] Ross, H.E. & Brodie, E.E. (1987): "Weber fractions for weight and mass as a function of stimulus intensity", *Quarterly Journal of Experimental Psychology 39A,* pp 77-88.

[160] Royal National Institute for the Blind (2000): "See it right", Peterborough: Royal National Institute for the Blind.

[161] Schiffman, H.R. (1995): "The skin, body and chemical senses", in R.L. Gregory and A.M. Colman (Eds.), *Sensation and perception*, London: Longman.

[162] Schrope, M. (2001): "Simply sensational", *New Scientist*, 2 June, pp. 30-33.

[163] Selcon, S.J., Taylor, R.M. & Shadrake, R.A. (1992): "Multimodal cockpit warnings: pictures, words, or both?", *Proceedings of the human factors society 36th annual meeting*. pp 57-61

[164] Sherrick, C.A. (1968): "Studies of apparent tactual movement", *The Skin Senses*, edited by D.R. Kenshalo. Springfield, C.C. Thomas, pp 331-345.

[165] Sherrick, C.A., Cholewiak, R.W. & Collins, A.A. (1990): "The localization of low- and high-frequency vibrotactile stimuli", *Journal of Acoustical Society of America, 88(1),* pp 169-178.

[166] Sinclair, R.J. & Burton, H. (1996): "Discrimination of vibrotactile frequencies in a delayed pair comparison task", *Perception & Psychophysics, 58(5),* pp 680-692.

[167] Somerson, P. (2001, March 2001): "Click and lick", ZDNet, available: http://www.zdnet.com/smartbusinessmag/stories/all/0,6605,2677285,00.html, (2001, 25 September).

[168] Spence, C. & Driver, J. (1997): "Cross-Modal Links in Attention Between Audition, Vision, and Touch: Implications for interface Design", *International Journal of Cognitive Ergonomics, 1 (4),* pp 351-373.

[169] Stevens, J.C. (1982): "Temperature can sharpen tactile acuity", *Perception & Psychophysics, 31(6),* pp 577-580.

[170] Stevens, J. C. (1991): "Thermal sensibility", in M. A. Heller and W. Schiff (Eds.), *The psychology of touch*, Hillsdale, NJ: Lawrence Erlbaum Associates.

[171] Stevens, J.C. & Cain, W.S. (1970): "Effort in isometric muscular contractions related to force level and duration", *Perception & Psychophysics, 8,* pp 240-244.

[172] Stevens, J.C. & Mack, J.D. (1959): "Scales of apparent force", *Journal of Experimental Psychology, 58,* pp 405-413.

[173] Sumikawa, D., Blattner, M. M. Joy, K. and Greenberg, R. (1986): "Guidelines for the syntactic design of audio cues in computer interfaces", *Proceedings of the Nineteenth Annual International Conference on System Sciences*, Honolulu, Hawaii.

[174] Tan, H.Z., Durlach, N.I., Beauregard, G.L. & Srinivasan, M.A. (1995): "Manual discrimination of compliance using active pinch grasp: the roles of force and work cues", *Perception & Psychophysics, 57,* pp 495-510.

[175] Teghtsoonian, R. & Teghtsoonian, M. (1970): "Two varieties of perceived length", *Perception & Psychophysics*, *8*, pp 389-392.

[176] Van Doren, C.L. (1990): "The effects of a surround on vibrotactile thresholds: evidence for spatial and temporal independence in the non-Pacinian I channel", *Journal of Acoustical Society of America, 87(6),* pp 2655-2661.

[177] Van Erp, J.B.F. & Van den Dobbelsteen, J.J. (1998): "On the design of tactile displays", Report TM-98-B012, Soesterberg, The Netherlands: TNO Human Factors Research Institute.

[178] Van Erp, J.B.F. & Van Veen, H.A.H.C. (2001): "Vibro-Tactile Information Presentation in Automobiles", in: Barber, C., Faint, M., Wall, S. & Wing, A.M. (Eds.). Eurohaptics 2001, pp. 99-104. Birmingham: University of Birmingham.

[179] Verrillo, R.T. (1962): "Investigation of some parameters of the cutaneous threshold for vibration", *Journal of Acoustical Society of America, 34(11)*, pp 1768-1773.

[180] Verrillo, R.T. (1963): "Effect of contactor area on the vibrotactile threshold", *Journal of Acoustical Society of America, 35(12)*, pp 1962-1966.

[181] Verrillo, R.T. (1965): "Temporal summation in vibrotactile sensitivity", *Journal of Acoustical Society of America, 37(5)*, pp 843-846.

[182] Verrillo, R.T. (1966): "Vibrotactile thresholds for hairy skin", *Journal of Experimental Psychology, 72(1)*, pp 47-50.

[183] Verrillo, R.T. & Gescheider, G.A. (1983): "Vibrotactile masking: Effects of one- and two-site stimulation" *Perception & Psychophysics, 33(4)*, pp 379-387.

[184] Verrillo, R.T. & Gescheider, G.A. (1975): "Enhancement and summation in the perception of two successive vibrotactile stimuli", *Perception & Psychophysics, 18(2)*, pp 128-136.

[185] Verrillo, R.T. & Smith, R.L. (1976): "Effect of stimulus duration on vibrotactile sensation magnitude", *Bulletin of Psychonomic Society, 8(2)*, pp 112-114.

[186] Walker, J.T. (1972): "Natural visual capture in bilateral length comparisons", *Perception & Psychophysics, 11*, pp 247-251.

[187] Walker, J.T. & Scott, K.J. (1981): "Auditory-vsual conflicts in the perceived duration of lights, tones, and gaps", *Journal of Experimental Psychology: Human Perception and Performance, 7*, pp 1327-1339.

[188] Warren, D.H. (1979): "Spatial localization under conflict conditions: Is there a single explanation?", *Perception, 8*, pp 323-337.

[189] Weinstein, S. (1968): "Intensive and extensive aspects of tactile sensitivity as a function of body-part, sex and laterality", in: *The Skin Senses*, edited by D.R. Kenshalo. Springfield, C.C. Thomas, pp 195-218.

[190] Weisenberger, J.M. & Craig, J.C. (1982): "A tactile metacontrast effect", *Perception & Psychophysics, 31(6)*, pp 530-536.

[191] Werkhoven, P.J. & Van Erp, J.B.F. (1998): "Perception of vibro-tactile asynchronies", Report TM-1998-B013, Soesterberg, The Netherlands: TNO Human Factors.

[192] White, C.T. & Cheatham, P.G. (1959): "Temporal numerosity: IV. A comparison of the major senses", *Journal of Experimental Psychology, 58*, pp 441 - 444.

[193] Wickens, C.D: "Processing resources in attention", in: R. Parasuraman & D.R. Davis (eds.): *Varieties in attention*. London: Academic, 1984, pp. 63-102.

[194] Wickens, C.D.: "Processing resources in attention", in: R. Parasuraman & D.R. Davis (eds.), *Varieties in attention*, London: Academic, 1984, pp. 63-102.

[195] Wogalter, M.S., Rashid, R., Clarke, M.J. & Kalsher, M. (1991): "Evaluating the behavioral effectiveness of a multimodal voice warning sign in a visually cluttered environment", *Proceedings of the human factors society 35th annual meeting*, pp 718-722.

[196] Chisholm, W., J. White, G. Vanderheiden, (2001): "Web Content Accessibility Guidelines 2.0", W3C Working Draft 24 August 2001, latest version available at http://www.w3.org/TR/WCAG20/

[197] ISO 13407: "Human-centred design processes for interactive systems".

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 [16].

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

auditory icon: auditory information which uses a real world, easily understood non-speech sound to communicate information, either in a unimodal message or as part of a multimodal message (e.g. the sound of something dropping into a metal container)

NOTE: This can be used instead of the visual change in the Trash icon for blind computer users, or in addition to the visual change, to create a multimodal icon for all users.

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

earcon: auditory information based on the use of abstract patterns of sounds known as musical motives

emoticons (or emotional symbols): graphical symbols that convey emotions or emphasize the communication by messages

ergonomics: branch of science that aims to learn about human abilities and limitations and then applies that knowledge to improve people's interaction with products, systems and environments

NOTE: Ergonomics is known as human factors in North America.

icon: within the field of human-computer interaction, graphic on a visual display terminal that represents a function of the computer system

NOTE: See [10].

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 [16].

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 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 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?

NOTE: For example, the navigation bar besides wordprocessors indicate 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.

pictogram: visually perceptible figure used to transmit information independently of language

NOTE 1: See [10].

NOTE 2: Also defined as "Graphical symbol which depicts objects or actions" [14]. The term "pictogram" is used for the graphical representation of a function or element of a user interface and includes both "icons" and "symbols" (see [2]).

symbol: graphic on a visual display terminal that provides an abstract representation of a function of the terminal or of the telecommunications network

NOTE: See [2] and [14].

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 [16].

3.2 Abbreviations

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

AT	Assistive Technology
ATM	Automatic Teller Machine
CSS	Cascading Style Sheets
DfA	Design for All
EU	European Union
GUI	Grafhical User Interface
HTML	HyperText Markup Language
ICT	Information and Communication Technologies
ICTSB	ICT Standards Board
IM	Instant Messaging
ISO	International Standards Organization
PC	Personal Computer
PDA	Personal Digital Assistant
SMS	Short Message Service
STF	Special Task Force
TC-HF	Technical Committee-Human Factors
URL	Universal Resource Locator
WAI	Web Accessibility Initiative
W3C	World Wide Web Consortium
WWW	World Wide Web
XML	eXtended Markup Language

4 An introduction to the Human Factors guidelines for multimodal symbols

Multimodal interfaces can achieve more natural and effective human-computer interaction by integrating a variety of signals by which humans usually convey information. The fact that the world is multimodal means that the recognition of a unimodal symbol may be hampered, even if it represents a real world equivalence but in isolation.

The explicit need to develop a multimodal interface may stem from several requirements. These requirements can, amongst others, be related to the user, to the task demands, to the character of the information, or to environmental constraints. Numerous specific user requirements have already been discussed in previous clauses. An example of a high task demand is reading an SMS message while driving. Presenting information to multiple modalities may release the overload on one of the channels, and will in general expand the information processing capacity of the user (see the multiple resource theory; Wickens, 1984 [194]). Also, environmental constraints may necessitate an additional modality. Examples include the limited visual displays of mobile devices, noisy environments, and the field of tension between public and private information presentation (e.g. a ringtone vs. a vibration to indicate an incoming call). A multimodal interface in these situations may be considered as a multi-unimodal interface. The effectiveness of a multi unimodal interface is based on redundancy, i.e. providing information (that is otherwise not available or in a degraded quality) to an alternative or additional information channel. This redundancy may be beneficial for all users and across all situations.

However, the advantage of multimodal interfaces lies not only in providing redundancy. A multimodal interface also allows to optimally allocate information to a specific channel, based on the match between the characteristics of the information and those of the sensory channel. It is therefore important to make an inventory of the strong and weak points of each sensory modality. Therefore, in clause 7.1, we start by presenting unimodal guidelines, ordered to the five human senses. In clause 7.2, we integrate the unimodal guidelines, discuss specific multimodal effects, and formulate the multimodal guidelines. It is important to note that the field of designing multimodal interfaces is young; the present clauses may therefore be tentative in some places.

5 Human Factors guidelines for multimodal symbols

5.1 Guidelines for unimodal symbols

Clause 7.1 discusses unimodal icons, symbols and pictograms per sensory modality: vision, audition, haptics, olfaction, and gustation, respectively. Each clause consists of an introduction, followed by the guidelines. Each guideline consists of four parts: a letter indicating the sensory modality (v, a, h, o, and g, respectively), a number, a title (all in bold and italics), and a description of the guideline (intended):

X [modality] # [number]. Title

Description [intended text body, describing the guideline with references when necessary].

5.1.1 Vision

5.1.1.1 Visual symbols, icons and pictograms: introduction

We will discern three basic forms for visual symbols: text, graphics and animation.

Text and graphics

Text is commonly used in all ICT products. The use of easily readable text will increase the number of both visually impaired and elderly people who can use an ICT product or service. The guidelines on the use of text are more or less clear, complete, and straightforward.

On the other hand, graphics in the form of icons, symbols and pictograms are also used extensively in ICT products and services to facilitate interaction between the product or service and the user. Graphics can provide a language-independent means of communicating information to the user. They are part of a graphical interface that can facilitate the user's ability to learn, understand and remember functions of the system, and aid in the manipulation of these elements.

Typically, a GUI draws on a user's environment to provide a metaphorical representation of the user's tasks. A metaphor provides an analogy to concepts already familiar to the user, from which the user can deduce the system's use and behaviour. Symbols can express the metaphor directly, as graphical representations of the metaphorical objects. They may also directly represent a physical object.

ISO/IEC 11581 [19] to [25] deals extensively with standards for graphics in computer-based products and is highly relevant in this context. It also presents a conceptual framework to relate the function within the system to the symbols used, which can be potentially be extended to multimodal symbols.

Animation

Examples of simple animations are the outline zoom which accompanies the opening of an icon (animated transition), and animated progress-bars (feedback). An example of a more complicated animation is the Ananova (2001) [33] news service, which features a synthetic, animated character as its anchor woman, and this has brought animated characters to the attention of the general public. Ananova is realistically rendered, has animated facial expression and some gestures (notably head movement) and synthetic speech output with lip synchronization. There are now several companies that deliver technology for the production of synthetic animated characters, e.g. Digimask (2001) [80] and Haptek (2001) [116]. Other implementations focus more on gestures (e.g. Microsoft® Agent), where gestures play an important role when applying these characters in tutorials.

Animation is a particularly useful component of the presentation of visual information as it draws the attention of the user to the information. Human beings are particularly good at monitoring their visual field for change, even if this is in the periphery of the visual field (round the edges) rather than in the direction we are looking. Thus, people will detect small motions at the edge of their vision, and instinctively look in that direction, a capability symbols designers may capitalize on. Of course, this can also be over-exploited, as too much motion can be very distracting and even lead to motion sickness. Ludi (2000) [142] argues that animated icons can increase the accessibility of graphical user interfaces for visually impaired users with some residual vision. Adding animation in addition to simple enlargement can make icons distinguishable from one another, even for visually impaired users.

Baecker, Small & Mander (1991) [34] describe the benefits of animation by asking a series of questions about a system or interface that animation can help answer: *Identification*: What is this?, *Transition*: Where have I come from and gone to?, *Orientation*: Where am I?, *Choice*: What can I do now?, *Demonstration*: What can I do with this?, *Explanation*: How do I do this?, *Feedback*: What is happening?, *History*: What have I done?, *Interpretation*: Why did that happen?, *Guidance*: What should I do now?

In desktop GUIs, symbols are used for most of the purposes listed above. Animated characters may take the role of animated icons for many of the uses outlined above, for instance, guidance, demonstration, and explanation. Orientation and transitions could possibly be made clearer through the use of animated gestures, and feedback could be enhanced through facial expressions and gestures. Synthetic animated characters are clearly different from single, animated icons. However, animated characters may take on the role that animations have in an animated icon and be used in conjunction with a static icon.

The set of guidelines on the design and use of animations is not as complete is on text and graphics. Some basic questions still need to be resolved. For example, Baecker et al. (1991) [34] identify the need for guidelines regarding length, content, ordering and visual representation which will guarantee effective animations. Dormann (1994) [82] suggests that the theory of design may offer more in the form of design guidance than traditional methods, and discusses principles of visual rhetoric in this context. However, the paper does not offer concrete guidelines or design strategies. Ludi (2000) [142] outlines a study to be performed where the aim is to find optimal sizes for animated icons for partially sighted users. However, the results from this planned study have not been reported yet.

5.1.1.2 Visual symbols, icons and pictograms: guidelines

Guidelines for text

If the following guidelines are not used as the default values for the presentation of text, there should be an easily accessible option to switch to these options. (Based on Royal National Institute for the Blind, 2000 [160]).

V1. Use a minimal typesize of 12 point

The size of the type significantly affects its legibility and is one of the most important features to bear in mind. A minimum of 12-point type should be used, although the use of 14 point will allow more people with sight problems to use a system.

V2. Use a high contrast

The contrast between the background and the type is also extremely important. The better the contrast, the more legible it is. The size and weight of the type will also affect contrast. If using white type, make sure the background colour is dark enough to provide sufficient contrast.

V3. Stick to familiar typefaces

The choice of typeface is less important than size and contrast. As a general rule, stick to typefaces that people are familiar with and will recognize easily. Avoid italic, simulated handwriting and ornate typefaces as these can be difficult to read.

V4. Avoid capitals

Avoid excessive use of capital letters, as they are generally harder to read. A word or two in capitals is fine but avoid the use of capitals for continuous text.

V5. Spacing between lines

The space between one line of type and the next (known as leading) is important. As a general rule, the space should be 1,5 to 2 times the space between words on a line.

V6. Avoid light type weights

People with sight problems often prefer bold or semi-bold weights to normal ones. Avoid light type weights.

V7. Make numbers easy discernible from letters

If numbers are important in the system, choose a typeface in which the numbers are clear. Readers with sight problems can easily misread 3, 5, 8 and 0.

V8. Combine text and pictures with caution

Avoid fitting text around images if this means that lines of text start in a different place, and are therefore difficult to find. Set text horizontally, as text set vertically is extremely difficult for a partially sighted reader to follow. Avoid setting text over images, for example photographs. This will affect the contrast and, if a partially sighted person is avoiding images, he or she will miss the text.

Guidelines for graphics

The general requirements and recommendations for visual graphics will be presented here, ISO/IEC 11581 [19] to [25] should be consulted for further information.

V9. Make icons highly discriminable

The appearance of an icon when representing one state or mode of the computer system should be clearly discriminable from that representing another state or mode.

V10. Retain comprehensibility throughout changes

An icon should remain comprehensible and discriminable through any changes in appearance due to changes in its state or mode, in the environment for which it was intended.

V11. Be cautious with overlapping icons

Whenever an icon is moved to a position overlapping another icon, but not in such a way as to activate any sensitive regions, the overlapping sensitive region of the moved icon should be on top of the other icon.

V12. Ask for appropriate user permission

Interacting with icons should not destroy any user data without user permission.

V13. Use more than colour coding alone

Colour should not serve as the only informative element to distinguish between icons.

V14. Use additional graphics for specific functions

Additional graphics may be incorporated into icons to indicate more specific functions, however the resulting icons should remain discriminable.

V15. Provide consistent visual appearance within icon sets

The visual appearance of icons should be consistent within the set of icons. This means that within one set, icons should be displayed using similar graphical style, for example, a similar degree of realism.

V16. Preserve comprehensibility and discriminability across displays

If the icons are presented in different styles on different displays, the design of the icon should take this into account to preserve comprehensibility and discriminability of the icon and its principal components.

V17. Preserve comprehensibility and discriminability across different aspect ratios

If the icons are used on various displays that cause the icons to be displayed at different aspect ratios, icon design solutions should take this into account to make the appearance of the icon as similar as possible to the intended graphic.

V18. Assure comprehension

All available icons should be comprehensible. When first time comprehension is not a usability requirement, then icons should be learnable and discriminable.

V19. Label icons consistently

The location of any user-modifiable label relative to the icon should be consistent within any environment or any collection of environments designed for use together.

V20. Prohibit interference with animation

Animation should not reduce the comprehensibility and recognizability of an icon.

Guidelines for animations

Through the iterative design and subsequent empirical evaluation, Baecker et al. (1991) [34] uncovered important design issues with respect to animated demonstration icons. These design issues point to preliminary guidelines and offer directions for further research.

V21. Animate within icon boundaries

Animate within the icon boundaries, so as not to occlude other icons or obscure the user's context.

V22. Keep animations simple

Discrepancies in the icon animation might throw a user off track with respect to its meaning. Animations need to be kept simple, both visually and conceptually in order to avoid misinterpretation and confusion.

V23. Animate serially

Running all demonstration animations simultaneously may be "too busy". Rather start the animation in the icon when the cursor passed into the icon's bounding box.

V24. Do not automatically animate

Well-designed static icons might still be easier to comprehend than animated ones because they are less busy and can more effectively capture the essence of the underlying referent.

5.1.2 Audition

5.1.2.1 Auditive symbols, icons and pictograms: introduction

The use of auditory information is rapidly becoming an important issue in the development of interfaces for ICT systems, both for unimodal information and for multimodal information. We discern the following two basic forms: speech output and non-speech output.

For non-speech audio, two different theoretical frameworks have been developed and will be outlined here. One approach was initially proposed by Gaver [91] and draws heavily on the "ecological" approach to the study of psychological phenomena developed by the psychologist J.J. Gibson (Gaver [92], [93]; Gibson [101]). This approach suggests that complex, real-world, immediately recognizable sounds, such as the sound of a door creaking or a bottle breaking should be used metaphorically in computer-based systems. The other approach, developed by Blattner and her colleagues (Blattner, Sumikawa and Greenberg [40]; Sumikawa, Blattner, Joy and Greenberg [173]) draws of the musical qualities of auditory information. This approach suggests that relatively simple tones can be combined in musical motives to provide auditory information, either as a unimodal system or as part of a multimodal system in ICT interfaces.

Gaver coined the term "auditory icon" to describe the use of real world sounds in interfaces in a manner analogous to visual icons and Blattner and her colleagues have used the term "earcon", starting from the idea that the word icon sounds like eye-con, and an auditory equivalent is therefore an ear-con. Some authors now use the term "auditory icon" to refer to the real world sounds proposed by Gaver's approach and "earcon" to the musical motives proposed by Blattner's approach, although this distinction is by no means universal, and other authors use the terms interchangeably. This review will use the terms to refer to the two distinct uses of sounds for ICT interfaces.

The ecological approach to the use of auditory information in ICTs

Gaver [91] initially proposed that sounds in computer (and by extension other ICT) interfaces should be "ecologically" meaningful in the same way that visual icons are meaningful. That is, they should provide an easily understood metaphoric link between a function in the ICT and an object or action in the interface to the ICT (see ISO/IEC 11581-1 [19]). Thus we understand the meaning of the Trash Can icon in the Macintosh GUI by drawing an analogy with our understanding of trash cans (or rubbish bins) in the real world as a place where things can be thrown away. Gaver suggested that we should extend the way we use sounds in the real world into the virtual computer environments of ICTs. Thus in real world listening we are interested in finding out information about the sources producing the sounds (i.e. objects, events and actions) and the environments in which they are produced, rather than the characteristics of the sounds themselves (which he suggests occurs more in musical listening). Thus the sound which an object makes when tapped tells us whether the object is hollow or solid and we may well tap it specially to obtain this information.

Gaver suggested two ways in which the virtual objects in ICT interfaces could also produce sounds, either unimodally or as part of a multimodal interaction, which can give useful information about those objects. The most effective auditory icons are those where the mapping from information to its representation is *nomic.* A nomic mapping is a necessary mapping caused by the physics of the situation in which it is produced. In the case of sounds, a given event necessarily produces a certain sound. So paper being riffled has a particular sound with which we are generally familiar and will never make a sound like water being poured out of a bottle; it is also practically impossible to riffle paper without making this sound. A second type of mapping between information and representation which Gaver proposed is where a metaphorical relationship is used. Metaphorical mappings make use of similarities between the item to be represented and the representing system. They are not wholly arbitrary, yet they do not depend on physical causation. Metaphorical mappings include structure mapping (Gentner [97]) in which similarities between the structures of the two systems are exploited.

Two problems arise with Gaver's approach. First, there may be many-to-many mappings even amongst nomic relationships which obscure the informational distinctiveness one wishes to convey with the sound when using them in an ICT interface. And second, although we may be familiar with the sounds of the objects in our real world environment, because most people usually see things as well as hearing them, we may not be able to recognize and use the information from these sounds without seeing their associated visual representation.

The musical pattern approach to the use of auditory information in ICTs

Blattner and her colleagues (Blattner, Sumikawa and Greenberg, 1989 [40]; Sumikawa, Blattner, Joy and Greenberg [173]) have proposed a very different approach to the use of auditory information in ICT interfaces, based on the use of musical patterns of sounds. In their system of "earcons", the basic building block is the *motive*, a single note of a particular pitch and duration, or a rhythmicized sequence of such notes. The simplest earcons are one element earcons, consisting of a single motive. Blattner et al. suggested that these simplest earcons should be used to represent basic, common computer entities such as error messages, system information, windows and files.

More complex earcons can be constructed using the principles of combination, inheritance and transformation. For example, if the action "create" is represented by a single note earcon and the object of type "text string" is represented by a two-note earcon, then these two earcons can be combined to form a compound earcon of three notes to represent the message "create text string". Blattner et al. also suggested that some of the inherent disadvantages of auditory icons can be avoided by using earcons. The problem of the many-to-many mappings between items to be represented and their natural sounds is avoided by creating arbitrary mappings. However, this means that the potential for the natural memorability of those mappings is lost. Designers should also test potential earcons for the memorability and the best mappings to objects, events and actions empirically, so that the set of earcons is optimal psychologically. Finally, auditory icons may have greater initial memorability, and be more appropriate for systems with little or no learning phase (e.g. walk up and use systems), whereas earcons may have greater long-term memorabilitiy and be more appropriate for more complex systems which users need to invest time in learning to use (e.g. computer applications such as word processors or frequently used telephony services).

Hankinson and Edwards [115] argued that using a musical grammar during the design of a set of earcons can impose a number of valuable constraints upon the choice of earcon motive. In a musical grammar, the basic units could be notes, chords, rhythms, pitch contours or even larger musical phrases. For example, when earcons representing a specific object and a specific action are played together the "melody" of each earcon (object or action) is designed such that objects and actions that are allowed in combination have the same basic chord, and therefore the "melodies" (i.e. the compound earcon) will sound harmonious. Objects and actions that are not allowed in combination are designed such that their basic chords clash and cause the "melodies" to sound discordant.

Example application: translating GUIs to auditory interfaces

There has been several experimental systems that attempt to translate graphical user interfaces (GUIs) into auditory interfaces, notably the Sonic Finder, The Sound Shark system, ARKola, EAR and ShareMon [94]. The Mercator system (Mynatt & Edwards [148]) is one of the most ambitious attempts, and tried to provide transparent access to X Windows applications for computer users who are blind or severely visually impaired. Mynatt & Edwards (1992) [148] point out some important design considerations, and among them to which degree the auditory interface should mimic the visual system. Two extreme approaches are outlined, one in which every aspect of the visual system is modelled in the auditory system for compatibility purposes (e.g. to support collaboration between sighted and non-sighted users). The other extreme approach is where the auditory interface is completely different and optimized for its medium.

Mynatt & Edwards [148] suggest a compromise, where some features that are related only to the graphical interface do not need to be modelled in the auditory interface, and where compatibility is sought after for the features that are modelled. In the Mercator, the interface is translated at the level of the interface components, such as menus, dialogue boxes and buttons. The Mercator interface tries to mimic some of the attributes of graphical interface objects, and uses auditory icons to convey the type of auditory interface components (AIC). To convey attributes of auditory interface components (e.g. selected, unavailable, has sub-menu), Mercator uses *filtears.* Filtears have the capability of presenting attributes across different types of interface objects. Table 5.1, extracted from Mynatt & Edwards [148] show some applications of filtears to convey auditory interface components attributes.

Table 5.1: Using filtears to convey AIC attributes. From Mynatt & Edwards (1992)

Attribute	Auditory Interface Component (AIC)	Filtear	Description
Selected	all buttons	animation	Produces a more lively sound by accenting frequency variations.
Unavailable	all buttons	muffled	A low pass filter produces a duller sounds
Has sub-menu	menu buttons	inflection	Adding an upward inflection at the end of an auditory icon suggests more information.
Relative location	lists, menus	pitch	Map frequency (pitch) to relative location (high to low).
Complexity	containers	pitch, reverberation	Map frequency and reverberation to complexity. Low to large, complex AICs and high to small, simple AICs

Although mapping GUIs to auditory interfaces deals with complete applications or interface environments, and not just symbols (as is the main topic of the present document), the strategies here are still relevant. Symbols are used to interact in applications and program environments, and in the design of symbols this needs to be taken into account, for instance by designing symbols such that interaction attributes (e.g. selected, unavailable, shortcut) can be added.

5.1.2.2 Auditive symbols, icons and pictograms: guidelines

Guidelines for speech output

A1. Use different voices for different interface elements

Different interface elements should be distinguishable by using different voices. For example, commands (links, forms, etc.) and information content (text, images).

A2. Respect the echoic store size when designing menus

Respect the limits of the human working memory and the echoic store (approx. 3-4 elements) when designing menus.

A3. Comfortable output speed is 160 words per minute

A user can listen to 150-160 words per minute comfortably. This speed should not be increased unless the user configures so (e.g. a blind user). Absolute speed output limit is 210 words per minute. User should be able to control the speech speed output.

A4. Present messages serially

Never present two verbal tasks at the same time (e.g. two messages). Every prompt must wait for a response.

A5. Speech and non-speech may be combined

A particular system can mix voice output with other non-speech sounds (e.g. earcons). For instance, to indicate that a download process is finishing, etc.

A6. Employ an adjustable number of control levels

There should be different control levels depending on experience. For example, more prompts for novel users, fewer for experts. Feedback should be relevant, never provide unnecessary feedback. An undo or back option is required, or otherwise, present a dialogue to confirm an action.

A7. Simulate human voices as much as possible

Modulate output as much as possible to human voice in synthetic speech systems. If it is possible, simulate accent, age, and sex of user. Also, use simulations which are similar to human prosody (tones related to the task, e.g. higher tone if it is a warning, interrogation, etc.)

A8. Eliminate non-relevant speech

All non-relevant verbal information has to be eliminated.

A9. Present a summary first

Before offering detail on a page of content, offer a summary or introduction of what is available in it. E.g. this is a page containing 5 images, 2 text paragraphs, etc.

A10. Present information in a logical order

Movements between interface elements must be made in a consistent and logical way: top-down, left to right.

A11. Preferably integrate speech output in a multimodal setting

Integration of speech output with other modalities is always desired. For instance, integrating a vocal interface with a touch display or with a Braille output device. Speech information may be very poor or difficult to use, so other mechanisms can greatly help to make it usable.

Guidelines for non-speech output

A12. Use a nomic mapping

Where possible, use nomic mappings between function in the ICT and the auditory icon, this is a necessary and sufficient mapping cause by the physics of the situation in which it is produced. If a nomic mapping is not possible, use a mapping involving a clear and easily understood metaphor.

A13. Real world sounds enlarge the recognizability

Use real world sounds for auditory icons that can be easily recognized even without accompanying visual information. This will need to be validated with user testing.

A14. Parameterization can increase the information content

Using Parameterization (varying one or more dimensions of the auditory icon) to increase the information content of the auditory icon. For example, if the auditory icon for a text document opening is a paper riffle sound, then the length of a paper riffling sound can be used to indicate the size of the document.

A15. The learnability of the auditory information should match the system

Auditory icons are good for ICT systems with no anticipated learning phase (e.g. walk up and use systems such as ticket machines) as they draw on the users' prior understanding of the sounds used. Earcons are good for ICT systems were users will invest time in learning the system (e.g. word processors, frequently used telephony services).

A16. Use multiple harmonics

Use musical instrument timbres, as simple tones such as sinewaves or square waves have been found ineffective (Brewster, 1994) [45]. Where possible, use timbres with multiple harmonics as this helps perception and can avoid masking. Timbres that are subjectively easy to tell apart should be used. For example, on a musical instrument synthesizer, use "brass" and "organ" rather than "brass1" and "brass2", as the former pair will be more distinct. However, instruments that sound different in real life may not sound as different when played on a synthesizer, so care should be taken when choosing timbres. Using multiple timbres per earcon may confer advantages when using compound sounds.

A17. Do not rely on absolute judgements

If listeners are to make absolute judgements of earcons then pitch/register should not be used as the distinguishing attribute, as these have been shown to be not effective by themselves (Barfield, Rosenberg and Levasseur [35]; Brewster, Wright and Edwards [50]). A combination of register and another parameter will give better rates of recall. If register alone must be used, then there should be large differences between earcons, and even then it may not be effective. Two or three octave differences give better recall. Much smaller differences can be used if relative judgements are to be made.

A18. Pitch structures can help to differentiate earcons

Complex intra-earcon pitch structures are effective in differentiating earcons if used along with rhythm or another parameter.

A19. The pitch should be within the hearing range

The maximum pitch used should be no higher than 5 kHz (four octaves above C3) and no lower than 125 Hz to 150 Hz (the octave of C4), so that the sounds are not easily masked and are within the hearing range of most listeners (Patterson [152]).

A20. Timbre and pitch must match

Care should be taken that the pitches used are possible and appropriate for the chosen synthesized timbre, as not all instruments play all pitches. For example, a violin may not sound good if played at very low frequencies. If a wide range of pitches is needed, then timbres such as organs or pianos are effective.

A21. Make rhythms as different as possible

Putting different numbers of notes in each rhythm has been shown to be very effective (Brewster [45]). Patterson [152] noted that sounds are likely to be confused if the rhythms are similar even if there are large spectral differences.

A22. Optimize the presentation rate

Earcons should be kept as short as possible so they can keep up with the interactions in the other modalities, particularly the visual modality. It may be possible to play two earcons in parallel to speed up presentation (Brewster, Wright and Edwards [49]). Earcons with up to six notes play in one second have been shown to be usable. In order to make each earcon sound like a complete rhythmic unit, the first note should be accented (played slightly longer) and the last note should be slightly longer (Handel [113]). Changing the tempo, that is speeding up or slowing down the sounds, is another effective method for differentiating earcons. Small note lengths might not be noticed, so do not use notes less than 0,0825 seconds in duration. However, if the earcon is very simple (one or two notes), then it has been shown that notes as short as 0,03 seconds can be used (Brewster, Wright, Dix and Edwards [48]).

A23. Keep the intensity comfortable

Great care should be taken over the use of intensity in the design of earcons as it is the main cause of annoyance due to sound (Berglund, Preis and Rankin [39]). The overall sound level should be under the control of the user (via a volume control mechanism). Earcons should be kept within a narrow intensity range so that if the user turns down the volume, no sound will be lost; if the user turns up the volume, then no one earcon should stand out and be annoying.

A24. Do not rely on Intensity coding alone

People are not good at making absolute intensity judgements (Buxton, Gaver and Bly [54]). Therefore, intensity should not be used on its own for differentiating earcons. If it must be used in this way, there should be large differences between the intensities used. This may lead to annoyance on the part of users, as it contravenes the previous guideline. Some suggested ranges are - minimum intensity: 10 dB above threshold; maximum intensity: 20 dB above threshold (Patterson [152]).

A25. Use spatialized sound to play earcons simultaneously

This may be two dimensional (stereo) position or full three dimensional position if appropriate hardware is used. This may be very useful for differentiating parallel earcons playing simultaneously (Brewster, Wright and Edwards [49]). It can also be used with serial earcons, for example each family of earcons might have a different location. The accuracy of both two and three dimensional spatial localization depends very much on the hardware used. Care needs to be taken in choosing locations that will be clearly differentiated in the target environments.

A26. Making earcons attention capturing by rhythm or pitch

In many cases designers want to their sounds to capture the listener's attention. This can be achieved in different ways. It can be done by using sufficient intensity. This is crude but effective (and very common). However, it is potentially annoying for the primary user and others nearby, so other methods are recommended. Rhythm or pitch can be used (perhaps combined with lower intensity), for example, because the human auditory system is very good at detecting dynamic stimuli. If a new sound is played, even at a low intensity, it is likely to grab a user's attention (but not that of someone nearby). Alternatively, if the rhythm of an earcon is changed (perhaps speed up the tempo or slowing it down), this will also capture attention. Other techniques for making sounds attention-capturing include: high pitch, a wide pitch range, rapid onset and offset times, irregular harmonics and atonal or arrhythmic sounds (for further information see Edworth, Loxley and Dennis [83]). The opposites of most of these techniques can be used to make sound avoidable, but in this case the main parameters are low intensity and regular rhythm.

A27. Put a gap between consecutive earcons

When playing serial earcons one after another, use a 0,1 second gap between them so that users can tell where one finishes and the next starts (Brewster, Wright and Edwards [50]).

A28. Earcons can provide good navigational cues

When designing the earcons for a menu hierarchy. Start with a neutral sound to represent the root of the hierarchy. Use a combination of instrument and register at the next level (Level 2) and then rhythm at the next (Level 3). These allow the structure to be expanded easily. If a new subtree is needed at level 2, then all that is necessary is a new instrument. If expansion is needed at Level 3 then more rhythms can be created. For deeper levels then the cues must be made very obvious (because users have much to remember to work out their location when they get deep into the hierarchy), for example a level indicator instrument might need to be added.

A29. Make the earcon robust for low quality reproduction

Low quality sound reproduction, as found in standard telephone systems, is adequate for differentiating between earcons. However, if earcons are to be used in these conditions (as opposed to in interfaces with high quality sound reproduction such as PCs), the earcons need to be designed to make the differences clear and appropriately evaluated to establish their usability in the target environments.

A30. Provide active training

Allowing users time to learn earcons ("active training") has been shown to be very beneficial. Providing both written instructions, and the opportunity to listen to the earcons has been found to significantly improve the recall of earcons (Brewster [46]). This could be provided by online tutorials (for PCs, PDAs etc) or free call time on a network (for fixed line and mobile telephony).

5.1.3 Haptics

5.1.3.1 Haptic symbols, icons and pictograms: introduction

When we think of interacting with objects by manipulating them, whether in the real or virtual (computer) world, we often think of touching and feeling them. Indeed, designers often discuss the "look and feel" of an ICT interface, even if this is currently largely an interface based on looks rather than its feel. However, unlike looking at objects which is conducted via the visual sense, manipulating objects, coming into direct contact with them, involves a rather complex set of sensory processes:

- Tactile or cutaneous perception - strictly speaking, refers to the contact of the skin with the outside world. The skin is sensitive to three primary qualities, mediated by a range - pressure (vibration), temperature and pain.

- Kinaesthesis or proprioception - as we move our limbs and body, we are aware of the position and movement of various parts of the body in relation to one another via feedback from receptors in the joints and the muscles. This gives us our kinaesthetic or proprioceptive sense.

- Vestibular sense - as we move our body through space, we are also aware of the position of our body in relation to the rest of the world, whether we are stationary or in motion, whether we are upright, leaning over or lying down. This is our vestibular sense (for further information on the sensory processes see Cholewiak and Collins [60]; Schiffmann [161]).

When we handle an object, typically both the tactile and kinaesthesis are relevant to the experience (often referred to as haptics; Heller [119]). The object exerts a certain pressure on our hands cutaneously which gives a sense of the weight and texture of the object and it also conveys a certain temperature to our hands and as we move our hands over the object our kinaesthetic sense gives us information about the size of the object. We discern three basic forms here: the vibro-tactile, the temperature, and the kinaesthetic sense.

The vibro-tactile sense

ETR 116 [30] defines a tactile display as the type of display that relies on the user perceiving the intended information by touch. An important distinction should be made between active and passive touch. Active touch is reserved for experiences in which the observer uses touch in an active and often exploratory way, such as in experiencing texture. This means that the display is often passive. Passive displays always present the same message, and the information transfer depends on activity of the user. Passive touch, on the other hand, requires no activity of the observer, but is based on an active tactile display. The most important class of those are the vibro-tactile displays, for instance, such as those used in mobile phones. An active display is able to present a variety of messages.

Most existing guidelines are related to passive tactile messages, such as Braille labels on controls, nibs on the keyboard and notches on smart cards (e.g. ETR 165 [5], ETS 300 767 [8], ES 201 381 [11]). However, technological developments around tactile display elements will soon lead to the possibilities of introducing active displays on a large scale. Nowadays, using vibration in mobile phones and pagers is common practise, just as the use of a vibrating computer mouse. These active tactile displays add the possibility of using the tactile sense as a full channel in the man-machine interface. The potential information transfer capacity of this channel is much larger than the 1-bit message "your phone is ringing". It may be foreseen that the use of the tactile modality will grow fast. Not only as an alternative information channel for people with special needs, but for all possible users. Especially in areas where the visual display possibilities are limited (e.g. in handheld devices), or where the auditive channel is unattractive (e.g. in mobile phones in public places). Furthermore, the tactile modality has several interesting characteristics. It is always ready to receive information (e.g. no need to "look" at the display), attract attention (which makes it suitable as warning device), can be intuitive (especially regarding spatial information (left is left, and right is right), and is private. At this moment, there is a lack of guidelines on the design and application of active tactile displays (with the exception of Braille displays and vibrating alerting devices), and on the interaction with the other sensory modalities (multimodal interaction). Since the pool of best practises is also very limited, the most important source to distil the guidelines from is the neurophysiological and psychophysical literature.

The skin is sensitive to numerous forms of energy: pressure, vibration, electric current, but also cold and warmth. In relation to display technology, by far the majority of the active tactile display is based on vibration. There are two major principles to generate vibration: electro-cutaneous (based on electrodes attached to the skin), and mechanical vibration. Although both techniques are quite different, psycho-physical experiments show that the characteristics of the skin are the same for both.

Although knowledge of the four different mechanoreceptors in the skin helps to interpret psychophysical results and the guidelines derived from it, presenting neurophysiological data is beyond the scope of the present document. Relevant overviews can be found in Kandel et al. [133], Johansson, Landstrom and Lundstrom [125], Bolanowski, Gescheider, Verillo and Chechosky [43], Bolanowski, Gescheider and Verillo [42], and Johansson [124]. Two points, however, are important. First, there is considerable spatio-temporal interaction within the tactile channel. And second, when a pure vibration is presented, the frequency determines which neurophysiological channel is activated. Often, one speaks of two frequency bands: the low frequency channel (frequency < 80 to 100 Hz), and the high frequency channel (frequency > 80 Hz to 100 Hz).

The temperature sense

One of the aspects of experience felt through the skin is temperature, the perception of warmth and cold. The human body has separate receptors for warmth and cold, hence different qualities of temperature can be coded primarily by the specific receptors activated. However, this specificity of neural activation has its limits. Cold receptors respond only to low temperatures but also to very high temperatures (above 45 degrees centigrade). Consequently, a very hot stimulus will activate both warm and cold receptors, which in turn evoke a hot sensation (see Stevens [170] for further detail).

Thermal information has been little used in interfaces to date, although a number of prototype systems have been developed. For example, MacLean and Roderick [143] developed an intelligent multimodal door interface, Aladdin, that incorporated auditory, haptic and temperature output. They noted that including the temperature output was the most experimental component of the display, since that medium has been less used than other haptic components in interface design. Although they developed a number of prototype scenarios for the use of the Aladdin system, these actually made little meaningful use of the temperature information. A prototype system has also added temperature information to the PHANToM haptic device (Ottensmeyer [150]), one of the high-end haptic devices used in research and applications such as telesurgery.

As such little use has been made of temperature information in multimodal systems thus far, no guidelines can yet be proposed.

The kinaesthetic sense

Devices that provide haptic output are currently becoming available for use in ICT systems. These started with very expensive and sophisticated devices such as the Impulse Engine from the Immersion Corporation (see http://www.immerse.com) and the Phantom (see http://www.sensable.com) which provide full three dimensional haptic or force feedback. However, recently two dimensional haptic feedback has been incorporated into computer mice and joysticks, and these provide an inexpensive and convenient method for incorporating haptic output into interfaces.

Two types of devices of particular current interest are the TouchSense-enabled mice from Logitech (the TouchSense technology was developed by the Immersion Corporation and has been licensed to a number of mouse and joystick manufacturers, see http://www.immersion.com/products/) and the ScreenRover from Betacom (see http://www.betacom.com/optic/screen.htm). Both these devices provide demonstrations of how force feedback information can be added to GUI environments by adding the sensation of borders to the edges of screen objects and textures to their surfaces. This haptic information can potentially be used in conjunction with a visual GUI or a screenreader to produce a audio-tactile multimodal system for visually impaired users.

A number of prototype implementations have added haptic information to the presentation of GUI systems (e.g. Miller and Zeleznik [146]). Unfortunately, little formal evaluation of user reactions to the haptic information in these systems is currently available. Prototype systems to make visual ICT systems accessible to blind people using a combination of auditory and haptic information have also been developed (Colwell, Petrie, Kornbrot, Hardwick and Furner [62]; Ramstein, Martial, Dufresne, Carignan, Chassé, and Mabilleau [156]) but these have not yet reached the marketplace.

Miller and Zeleznik [147] have proposed that haptic information is most suitable for use in several contexts in ICT systems:

- ***Anticipation***: Haptic information can be used to provide a "breakable" force (meaning the user can break through the force) resisting the user's motion and indicating the imminence of a qualitative change in the user's input before the user actually makes it, so that the user has the opportunity to back off from that change if it is undesired.

- ***Follow-through***: Haptic information can also be used to let the user know than an attempted qualitative change has actually been accomplished, so that they have the opportunity to correct their motion if they do not get this feedback.

- ***Indication***: Haptic information can provide an indication that a continuing qualitative condition remains in effect, possibly with quantitative information about the condition. For instance, a joystick with springs to return it to the centre position lets the user know that the control is away from the neutral position and also in what direction it is away from that position and possibly how far.

- ***Guidance***: Haptic information can be used to adapt the user's input with a bias towards some set of possible inputs. For instance, a straight edge is an example of guidance because it can be used to guide the user towards those points along a straight line at a particular position. As an example, consider a hypothetical radio station tuning dial with detents at user-settable positions (e.g. corresponding to favourite stations).

- ***Distinguishing directions***: Haptic information can be used to allow the user to make clear distinctions between locally orthogonal directions. This can be used to map different (but possibly related) controls onto different dimensions of the same input mechanism. One example from the real world is a flight yoke that allows rotations of a control about a shaft and translation of the control along the shaft.

Most research on kinaesthetic perception has focussed on the perceptions of exerted force, limb position and limb movement. However, the kinaesthetic system also uses the signals about force, position, and movement to derive information about other mechanical properties of objects in the environment, such as stiffness and viscosity (Jones, Hunter & Lafontaine [127]). An understanding of the perceptual resolution of the kinaesthetic system for such object properties is of particular importance to the design of haptic interfaces. Therefore, we give a concise overview of the results of studies on psychophysical scaling and JNDs for several parameters. The exponent of the power law (percept = constant * physical magnitude ^ exponent) is 1,7 averaged across studies (Stevens and Mack [172]; Eisler [85], [86]; Stevens and Cain [171]; Cain and Stevens [55]; Cooper et al. [67]), although some studies report an exponent smaller than 1 (Cooper et al. [67]; Jones & Hunter [128]).

The subjective level of force increases with time (Stevens & Cain [171]; Cain & Stevens [55]; Cain [56]). The JND for force is about 7 % (Jones [131]; Pang, Tan and Durlach [151]; Tan et al. [174]). The JND for stiffness (stiffness (the change in force divided by the change in distance) is much higher. It is difficult to present a general value for the JND of stiffness, since the different studies revealed considerably different JNDs. The JNDs reported vary between 19 % and 99 % (Jones &Hunter [129]; Roland & Ladegaard-Pedersen [157]). The JND values for viscosity (a change in force divided by a change in velocity, expressed in Ns/m) depend on the reference values. For small values, the JNDs are high: 83 % at 2 Ns/m to 48 % at 16 Ns/m (Jones and Hunter [130]). For higher values, the JND is lower. Reported values range from 9,5 to 34 % (Jones & Hunter [130]; Jones et al. [127]; Beauregard, Tan & Durlach [37]; Beauregard & Srinivasan [36]). Finally, the reported JNDs for mass (defined as the ratio of applied force to achieved acceleration) are relative uniform across studies: 10 % is found for weights of 50 g, and a smaller JND for weights above 100 g (Ross and Reschke [158]; Brodie and Ross [52]); Brodie [53]; Ross and Brodie [159]; Darwood, Repperger and Goodyear [77]; Hellström [120]). For very heavy weights, the JND decreases to 4 % (Carlson, Drury & Webber [58]).

5.1.3.2 Haptic symbols, icons and pictograms: guidelines

Guidelines on the vibro-tactile sense

The number of available guidelines on vibro-tactile sense and the number of evaluated applications is still very limited. Most of the guidelines below are based on basic research, haphazard statements in the literature, and an overview by Van Erp and Dobbelsteen [177].

H1. Optimize the detection threshold by frequency and location

Lowering the detection threshold may have advantages (e.g. regarding power consumption). When this is an important design consideration, lowest thresholds are found on glabrous skin with vibration frequencies around 250 Hz. Other factors that affect threshold are the form of the contactor, the presence of a rigid surround, and the indentation. The waveform affects the threshold and the perception. Although usually a sine is used as vibratory input, a square wave is most intense, sine is the smoothest, and triangle is in between.

H2. Subjective magnitude can be used to encode information

Encoding information by using different levels of intensity levels is possible, although the number of levels is small: not more than 4 different levels should be used between the detection threshold and the pain/comfort threshold. Psychophysical parameters indicate that there are two ways of enlarging the subjective magnitude of a stimulus: 1. enlarging the intensity for intensities near the threshold, and 2. enlarging the area of stimulation.

H3. Frequency can be used to encode information

No more than 9 different levels of frequency should be used for coding information. Difference between the levels should be at least 20 %. Presented with the same amplitude, the levels will also lead to different subjective magnitudes.

H4. Use an intensity between 15-20 dB above threshold

Comfortable stimuli range 15 db to 20 dB above threshold. Analogue to vision and audition, a high intensity tactile stimulus may lead to discomfort and finally to a sensation of pain. Amplitudes above 0,6 mm to 0,8 mm will result in a pain sensation.

H5. Use a frequency between 10 and 600 Hz

The human tactile channel is only sensitive to frequencies between 10 Hz and 600 Hz. However, on the outer sides of the range the thresholds are high; preferably, only frequencies between 50 Hz and 400 Hz should be used. Lowest thresholds are around 250 Hz.

H6. Use distal body parts if a high spatial resolution is required

Spatial location may be an important parameter in the design of tactile displays. When high acuity is needed, only the distal body parts will suffice (e.g. the fingers). When spatial acuities as low as 4 cm are acceptable, any locus will suffice.

H7. Apparent position can enlarge the spatial resolution

Although the phenomenon of apparent location may raise the number of distinguishable locations, the usage is questionable when the density of actuators is close to the spatial acuity.

H8. Higher resolution can be allowed for trained users

When the display is designed for trained users, the density may be higher.

H9. Spatial summation can enhance signal detectability

Spatial summation results in decreased thresholds if the contactor area increases, or if two (widely spread) areas are stimulated simultaneously with similar frequencies. If there is an advantage of using the mechanism of spatial summation (e.g. for enhancing delectability of weak signals), the stimuli should preferably be presented to hairy skin or with frequencies above 80 Hz. Furthermore, when stimuli are used which can provoke spatial summation, one should consider the fact that spatial summation may occur as an unwanted effect as well, for example when using simultaneous stimulation at different loci.

H10. Be aware of spatial masking

Spatial masking may occur when stimuli overlap in time, but not in location. Especially when using pattern recognition, the designer should be aware of the negative effects of masking. Both the detection and the identification of stimuli may be degraded. Using different channels may prevent masking.

H11. The time between consecutive signals of a temporal pattern should be at least 10 ms

When using a single actuator of a tactile display to encode information in some kind of temporal pattern (called touchtones by Van Erp and colleagues, see Schrope [162]), the time between signals must be at least 10 ms. Depending on the type of actuator and the load, a vibratory stimulus will take time to reach the set frequency, and may smother slowly. This is important for the temporal aspects of the presented stimulation.

H12. Temporal enhancement can affect the subjective magnitude of separated stimuli

Temporal enhancement of a second stimulus (resulting in higher subjective magnitude) occurs when two stimuli are separated by 100 ms to 500 ms. Temporal enhancement occurs only when the stimuli are in the same frequency band.

H13. Detection thresholds can be lowered by using temporal summation

Temporal summation results in lower detection thresholds when stimulus duration increases. The mechanism of temporal summation might be useful when there is a need to lower the detection threshold, although it imposes restrictions on the possibilities of using time-based or frequency-based manipulations, and it requires a minimum area of stimulation and a frequency above 60 Hz.

H14. Temporal masking can distort the perception of multiple stimuli

Masking can occur when two stimuli are presented to the same location, and when the onset of the target stimulus is within a certain time interval from the onset of a "distractor"; this interval ranges from -100 ms up to +1 200 ms.

H15. Prohibiting temporal masking by locus

Presenting stimuli to different loci prohibits temporal masking.

H16. Prohibiting temporal masking by frequency

Presenting temporal patterns to different neurophysiological channels (i.e. using low and high frequencies) prohibits temporal masking.

H17. Beware of adaptation

Adaptation corresponds to a change in the perception of a stimulus after prolonged stimulation Adaptation decreases the absolute threshold and the subjective magnitude. This is a gradual process that takes up to 25 minutes. The effect on the threshold is larger (up to 20 dB) than on the subjective magnitude (up to 7 dB).

H18. Recovery from adaptation

Recovery time is about half the adaptation time, and is faster for the subjective magnitude than for the absolute threshold.

H19. Preventing adaptation by using frequency

Adaptation can be prevented by using different neurophysiological channels (frequencies).

H20. Be aware of the occurrence of perceptual illusions

Stimuli that are presented closely in time and space can alter the percept and may even result in a completely new precept.

H21. Apparent location

Apparent location is the percent of a single stimulus induced by the simultaneous activation of two stimuli to different locations. The apparent location is in between the two stimulus locations and depends on the relative magnitude. Both stimuli should be in phase to evoke a stable apparent location. Apparent location may be used to increase the number of stimulus sites, without enlarging the number of actuators.

H22. Apparent motion

Presenting two or more stimuli in a specific spatio-temporal pattern can evoke apparent motion. Apparent motion can be used to simulate actual motion in for example tracking displays. When employing the mechanism of apparent motion. Most important parameters are the bursts duration (minimum duration of 20 ms) and the time intervals between the onsets of the consecutive stimuli.

H23. Vibro-tactile actuators can cause pain

Electrodes and vibrators can generate sufficient heat to cause painful sensation of heat as well as burns. For example, 12-mm-diameter vibro-tactile transducers are uncomfortable warm at continuous average power levels above only 62 mW (55 mW/cm^2). This lower power is likely due to the fact that the transducers (and mounting hardware) are in physical contact with the skin.

H24. The comfort of a tactile display is dependent on the method used to hold the display to the skin

The mechanical comfort of any tactile display is heavily influenced by the method used to hold the display to the skin. It is important to effectively maintain a fixed contact between stimulation source and the skin for adequate transfer of information. This problem has made most of the displays inconvenient and impractical to use. A compromise must often be made between performance and comfort. Care must be taken to minimize post stimulation skin irritation.

H25. The user must be able to adjustable stimulus intensity

There is a high variation in thresholds of sensation and pain between subjects, and over the lifespan (spatial and temporal acuity degrades with ageing). The stimulus level must preferably be adjustable by each individual user.

H26. Ensure comfort over longer periods of time

Tactile displays that are worn on the body, must be unobtrusive, and comfortable for longer periods of usage. The vibrations of the hand-arm should be limited. The most critical frequencies are around 12 Hz, the critical range is from 1 m/s^2 to 5 m/s^2.

H27. Prevent the emission of acoustic energy

The system must be limited in its emission of acoustic energy. Unwanted acoustic output may be a source of interference to persons or equipment near the display user.

H28. Avoid the spreading of vibration among vibrators

Especially when a nearby actuator vibrates at the same resonance frequency, there is a risk of passing vibration onto not activated vibrators. A rigid surround can reduce the spreading.

H29. Tactile stimuli can annoy the user

Do not annoy the user; tactile stimuli are hard to ignore if the user does not want to use them.

H30. Prevent interactions with manual control tasks

Be aware that vibrations do not interfere with a manual control task (vibration as tremor).

H31. Make tactile messages preferably self-explaining

Most people are unfamiliar with tactile signals in human computer interaction. This means that the tactile messages must preferably be self-explaining (in analogy one can speak of vibrocons, e.g. see Van Erp & Van Veen [178]). It also means that users will not experience tactile signals continuously, and have limited opportunities to learn to know the meaning of tactile messages (tactile continuity is low), like people learn to know the meaning of visual symbols.

H32. Mimic the real world

Keep in mind the touch experiences in the real world. Touch is for example used to perceive: mass, size, structure, resistance, pressure, orientation (edges, etc.).

H33.Composing multiple messages

Complex tactile messages must preferably be composed of well known, meaningful components. However, combing different vibro-tactile signals may alter the percept, for instance as the sum of two waves that are out of phase.

Guidelines on the kinaesthetic sense

On the basis of their work on haptic information for textures and objects in virtual reality and ICT interfaces, Furner, Petrie and colleagues developed the following guidelines (see Furner, Hardwick, Colwell, Bruns, Petrie, and Kornbrot [90]; Jansson, Petrie, Colwell, Kornbrot, Fänger, König, Billberger, Hardwick, and Furner [123]; Penn, Petrie, Colwell, Kornbrot, Furner, and Hardwick [153]).

H34. Users need to be able to easily discriminate between different simulated textures

Do not assume that physical variations in roughness are easily detected or discriminated from one another. Users may vary in their perception of texture, both in the size of the differences which they can detect and in the way they feel textures (e.g. what is rougher, what is smoother).

H35. The size of a virtual object may need to differentiate from the real world dimensions

Users perceive the sizes of larger virtual objects more accurately than those of smaller virtual objects. Users feel virtual objects to be bigger from the inside and smaller from the outside. The former observations both suggest that if it is important for users to perceive size accurately, virtual objects may need to deviate from their real world dimensions in the virtual world.

H36. Virtual objects need not follow the laws of physics

Virtual objects need not follow the same laws of physics as real objects, e.g. users can push through the surface of an object. Current technological constraints mean that virtual objects may not be able to simulate all aspects of their real world equivalents. This does not appear to disturb users greatly in terms of pushing through the surfaces of objects, but care should be taken if the laws of physics are broken in other ways.

H37. Provide exploring strategies

Users may need to learn strategies on how to explore virtual objects with a particular device. This is probably not time-consuming, but useful strategies should be provided for users.

H38. Do not use only haptics for complex objects

Users may not understand complex objects from purely haptic information; multimedia information may be required to give a sense of complex objects and what they mean. Complex objects that are made up of component objects may have very small spaces between the components into which the haptic pointer may slip. Users may have to remove the pointer from the gap in order to continue to explore the object. They may also be confused by objects by finding unexpected gaps in objects.

H39. Provide navigation information

Users may become "lost in haptic space". Provide navigational information support to try to avoid this problem. Users may have differing mental models of where the virtual space is and what part of the device is "touching" a virtual object. Watch for any consequences of these factors.

H40. Intensity differences to encode information should be dependent on the physical entity

Intensity differences when encoding information should be at least 10 % for force and mass, and 100 % for stiffness and viscosity.

5.1.4 Olfaction

5.1.4.1 Olfactory symbols, icons and pictograms: introduction

Current user interfaces mainly apply three senses: vision (e.g. text, graphics, and animation), audition (e.g. speech and non-speech sounds) and haptics (e.g. feedback in keyboards and mice). Humans have much richer capabilities than that, and olfactory displays are now seen as a possibility for inclusion in user interfaces.

Smell and the olfactory system have been researched extensively, and for different purposes, mainly in the food and perfume industries. The entertainment industry has also experimented with synthetic smell production, in the form of accompanying smells to enhance the experience of films (Lefcowitz [141], Somerson [167]). In the Aroma Rama and the Smell-o-vision systems, smells were released in cinema theatres in certain scenes of the film. In the John Waters film "Polyester" in 1981, the audiences were given "scratch and sniff" cards and asked to release smell at certain places during the film. These experimental systems were mainly novelties and not very successful, with reactions from the audiences reaching from allergic reactions to nausea.

The systems discussed above were all manually controlled, and the scents were all pre-produced. With respect to the inclusion of smell in the user interface, it only becomes interesting when the production of smell can be computer controlled and can be produced based on a computerized descriptions of particular smells. Then it will be possible to include olfactory displays in computer systems. For smell to gain acceptance among audiences there are many more factors that need to be in place, such as natural smelling odours, non-allergenic smells, etc.

The main idea of how an olfactory display would work is that the user has a peripheral device for smell production. This device is connected to the computer, and controlled by the computer. Using codified descriptions of smell, the computer can signal the release of a particular smell. A specific smell is generated by mixing a set of primary odours, most likely in the form of oil-based fragrances (Bonsor [44], Cook [64]).

There are (at least) two companies that have designed systems based on this technique. The peripheral device has a cartridge (that can be refilled or replaced) that holds the oils that produces the smells. DigiScents (see Bibliography), being the most high-profiled company, identified, coded and digitized thousands of smells and identified 128 primary odours that could be mixed to generate other smells (Bonsor [44]). The iSmell Personal Scent Synthesizer was given a lot of attention in the media, but early in 2001 the DigiScents company went out of business, without delivering any technology.

The SENX Scent device from the company TriSenx used a similar principle to produce scents, and got as far as putting a price on and taking pre-orders for their SENX machine and the SenxWare Scent Design Studio Software. However, the TriSenx web site is currently not active, and their sales department has not replied to email enquiries. It can be assumed that the company has encountered similar problems to DigiScents. It can therefore be assumed that as of yet there are no computer-controlled scent devices that are commercially available.

However, the principle of computer-controlled scent production has been demonstrated, in an early device delivered by TriSenx and also by "Jofish" Kaye at the MIT Media Lab. Kaye [134] has written an MSc thesis on Symbolic Olfactory Display, and has also built a prototype system for generating and dispensing computer controlled smell. One demonstration dispensed mint odour when the Nasdaq index went up, and lemon when it went down.

Kaye (2001) [134] discusses olfactory iconography, and introduces the two terms "olfactory icons" and "smicons". An "olfactory icon" is defined as scent output to convey information, where the scent is environmental and semantically related to the information to be conveyed (e.g. releasing gunpowder smell when a shotgun is fired in the computer game Quake). This is equivalent to the use of the term "auditory icon" when using auditive information in user interfaces.

Kaye defines "smicon" to be scent that has only an abstract relationship with the data it expresses (e.g. setting an olfactory alarm to release the scent of wintergreen at noon each day). This would be equivalent to earcons in the sonic world.

Kaye identifies some key issues relating to the use of smell that should be researched further, and these may also serve as preliminary guidelines for the use of smell in user interfaces.

5.1.4.2 Olfactory symbols, icons and pictograms: guidelines

O1. Olfactory displays should rely on users distinguishing different smells, not the strength or duration of smells

Humans are better at perceiving a change in a smell rather than the intensity of a smell. Kaye concludes that an olfactory display should rely on the user distinguishing multiple odour qualities, and not quantities. Kay points out that such a qualitative display makes it difficult to present clear sequences or progressions (e.g. that one scent is "more" than another). Kaye's system "Dollars & Scents", which emits a mint odour when the Nasdaq goes up, and a lemon odour when it goes down, is an example of using smell in a qualitative way.

O2. Smell is generally appropriate for slow-moving, medium-duration data

Smell lingers, and the duration of a smell may of course vary due to variations in air supply, ventilation, etc. You may introduce a new smell at the same time, or another instance of the same smell, but it may not be perceived as that.

O3. Parallel presentation may result in new percept

Two smells introduced at the same time will not necessarily be interpreted as the two distinct smells, but rather as a mixture, i.e. a different smell all together.

O4. Olfactory displays must take into account potential allergy and nausea reactions, or other discomfort in users

Smells can induce allergy and nausea in users, and this must be taken into account both in the use of specific odours and the general application of olfactory displays. Also, given that smells have a strong memory enhancing effect, some people may experience psychological discomfort when exposed to specific smells. In some cases, causing discomfort in users may be the effect that you want, as described in Kahn [132]. However, for general-purpose user interfaces, that is most likely not the case.

O5. Olfactory displays are appropriate as ambient displays

Smell is intrinsically an ambient and peripheral medium, and as such appropriate for displaying ambient information, i.e. information that is conveyed in our periphery and is processed in the background. This could for instance be presence information (e.g. that someone is present in a virtual room), or other background information about the current state of the system.

O6. Olfactory displays should take individual differences into account

It is well known that many different factors may affect our sense of smell temporarily (e.g. a common cold) or more permanently (e.g. due to smoking). Smells may also be interpreted differently and associated with different things, partly depending on cultural background: what smells "pleasant" for some, may seem "unpleasant" to others. This must be taken into account when using olfactory displays.

5.1.5 Gustation

5.1.5.1 Gustatory symbols, icons and pictograms: introduction

The basic receptors for taste in human beings, called taste buds, are specialized structures located in microscopically small pits and grooves throughout the oral cavity. On average, people possess about 10,000 taste buds, which are generally found in clusters lying within small, but visible elevations on the tongue, called papillae. Not all papillae are equally responsible to the four basic or primary tastes of sweet, sour, salty and bitter. That is, different regions of the tongue are more sensitive to specific taste stimuli than are others. The front of the tongue is most sensitive to the sweet taste, the back sides for sour, the front and sides of the tongue are most sensitive to the salt taste and the front (and especially the soft palate) is most sensitive to bitter (Schiffman [161]).

Gustation is a sense that has as of yet not been explored to any extent for use in user interfaces. Bonsor [44] refers to the patented technology of the TriSenx company (see also in clause 7.1.4 Olfaction), which allows users to "print" smells onto thick fibre paper sheets and taste specific flavours by licking the paper coated with the smell. Intended applications are customer sampling of product flavours before buying over the Internet.

The TriSenx company (as mentioned before) seems to be in hibernation, or even, out of business. It can be assumed that there are no commercially available computer-controlled "gustatory displays". A great deal of research and development is needed before such displays become technologically viable and useful as information displays.

5.1.5.2 Gustatory symbols, icons and pictograms: guidelines

G1. Take into account the factors that affect the consistency of the percepts

Taste sensitivity for a substance is affected not only by its chemical composition but also by a wide variety of stimulus and human variables, such as its concentration, the area of its application, the age and prior dietary conditions of the taster and the temperature of the substance.

G2. Cross-modal interaction affects the percept

The sense of taste also shows strong cross-modal interactions with the sense of smell, the cutaneous experience of the substance in the mouth and the visual appearance of the substance being tasted. If the sense of smell is reduced or eliminated by blockage of the air passages of the nose, such as by tightly pinching the nostrils or involuntarily when one has a cold or flu, two quite different food substances may taste surprizingly similar. For example, under such conditions of reduced smell, raw potato does not taste very different from apple. Such cross modal processing can be demonstrated even more dramatically by taking a number of fruits such as banana, apple and pear, pureeing them, colouring them blue with tasteless food dye and then asking people to taste them. Even with the appropriate smell sensations intact, with the normal texture and colour of the foods removed, people find it impossible to distinguish between these three very different fruits and often remake that they have no taste at all.

G3. Gustatory displays should take individual differences into account

Olfaction and gustation being highly related senses, most of the cautions attached to the use of olfactory displays in user interfaces also hold for "gustatory displays". What is a "good" taste for some will be a "bad" taste for others. Somerson [167] will certainly not be the only sceptic with respect to the use of such technology. Humans may also have temporary or more permanent taste impairments that make it difficult to use gustatory displays for computer systems.

5.2 Guidelines for multimodal symbols

Effective multimodal interaction requires that information presented to the different sensory channels is co-ordinated and made congruent informational as well as spatially and temporally (Kolers & Brewster [137]). Although stimuli may objectively be synchronized in time and space, this does not necessarily lead to a congruent percept across modalities. Numerous examples show how precept are affected across modalities. For example, in the spatial domain, vision dominates touch (sometimes called visual capture, Hay, Pick & Ikeda [118]), and found in e.g. estimating length and in perceived size (Teghtsoonian & Teghtsoonian [175]; Walker [186]). Warren and colleagues (see Freides [89]) found that vision biased proprioception 60 %, whereas proprioception biased vision approximately 35 %. Also in the spatial domain, touch dominates hearing (Lederman [140]; Pick, Warren & Hay [155]; Warren [188]). In the temporal domain incongruancies are also present, for example, the perceived duration of a sound is longer than that of a light of equal length (Behar & Bevan [38]; Goldstone & Goldfarb [105]), and intervals bounded by light flashes appear shorter than those bounded by brief auditory stimuli (Goldstone & Lhamon [106]; Walker & Scott [187]). A similar incongruency is present in the perception of visual and tactile time intervals (Werkhoven & Van Erp [191]). These examples indicate that sensory congruency is a non-trivial aspect of integrating sensory modalities. Knowledge of this congruency (or incongruency) is a prerequisite for the success of multimodal interfaces.

When combining and integrating different modalities, effects may be present that are not relevant in unimodal user interfaces. These effects are the main topic of this part of clause 7. The main outline is as follows:

1) The pros and cons of the different modalities are discussed. This might help the designer in allocating information to a specific sensory channel.

2) Second, specific multimodal effects are discussed (including sensory congruency). These effects are specific to combinations of modalities and are not present in unimodal information processing. They are therefore a potential pitfall.

3) Third, we discus examples and lessons learned of multimodal interfaces. Finally, we present the guidelines. Again, these guidelines consist of a letter (M for multimodal), a number, a title, and a description.

5.2.1 Allocating information to a sensory modality

The use of multi-unimodal information is one of the key solutions in a DfA approach to ICT design. Many of the needs of people with sensory disabilities, cognitive disabilities, and elderly people can be met within mainstream products if they are able to provide symbols and other output information in a range of different modalities. For example, the needs of blind and dyslexic people can be met if text and graphic information is also provided in synthetic speech output and the needs of hearing impaired people can be met if speech and sound information is also provided in text.

In approaching the use of multimodal information, designers may rightly wish to consider how to optimize the presentation of information in the different available modalities – if the choice is visual, auditory and haptic, which modality will be most appropriate for which component of the information? However, from a DfA perspective, designers should also consider a presentation method which involves considerable redundancy of information and allows users to adapt the presentation to fit their own requirements – whether these are permanent requirements (e.g. a particular sight condition which means they cannot use detailed visual information) or a temporary requirements (e.g. I'm driving my car at the moment and cannot look at the screen). Thus, for users with full capabilities in all modalities, a set of symbols which optimize the presentation in each modality can be designed. However, but for users with no visual capabilities (permanently or temporarily) the symbols can be presented using only auditory and haptic information, even if this is not the most optimal presentation combination.

In this clause, we summarize the pros and cons of the different sensory channels in table 5.2. Please note that this table presents an overview at a very general level. The match is always dependent on factors such as the technology available, and the table is therefore by no means absolute. For example, the technology on olfactory displays is in an early developmental state. The use of the table may help the designer to be aware of the different characteristics present **within** a single symbol, and help the designer to allocate each part of the information to the preferred sense (s). An example hereoff is a warning signal, which consists of a perception stage, possibly outside the primary area of interest (allocated to the auditive or tactile channel), a understanding stage and an action stage (preferably allocated to the visual or auditive channel). This methodology may indeed be an important step in developing a multimodal interface that is better than the sum of multiple unimodal interfaces.

Table 5.2: Match between the characteristics of the information and the sensory channel

characteristics of the information	vision	audition	Haptics	olfaction	gustation	references and footnotes
the information is time related (e.g. information that represents duration, interval, synchronization, or rhythm)	□	++	+	- -		Kirman [135]
the information is spatial (e.g. information that represents size or distance, e.g. the block on the scrollbar: the location and size provide information on the position in the document and the size of the document)	++	□	□			
2 D localization (the absolute and relative location in one or two dimensions, e.g. locating the trashcan on a desktop)	++	+	+			
3D localization (absolute and relative location in three dimensions)	[a]	+	+			[a] a 2D display requires the compression of one or more dimensions
the information has no world equivalent (Including abstract or coded information, e.g. a cancel button or a hyperlink but, but also engine rpm)	++	++ [a]	□			[a] synthesized speech
the information is private (information is intended to be perceived by a specific user or set of users only)	□	-	++			
the information is outside the primary area of interest or outside the area of spatial attention	- [a]	++	+ [b]	++		[a] the field of view of the user is limited [b] e.g. vibration function on mobile phones or other wearable devices
the information requires an optimized reaction time (under optimal perception)	++	+	□			
warning/alert (combination of perception, understanding and action)	- [a]	+	++ [b]			[a] primarily because of the limited field of view, employing peripheral vision requires motion. [b] should be a wearable
the information represents changes over time (most common case is the representation of a processes, e.g. downloading)	++	+	+/?	- -		
cultural generality (is there a general meaning of the symbol across cultures)	-	+	+/?	- -	- -	
memorability (i.e. the ease of recognition and identification of a former perceived symbol later in time)	+	++	-/?	++	+/?	

characteristics of the information	vision	audition	Haptics	olfaction	gustation	references and footnotes
the information represents a real world physical object [a]	++	-	□			[a] there is a near inverse relation with abstract information
the information should be persistence [a] (i.e. the information is available after initial presentation)	++	- -	++			[a] scores are reversed if the presentation should be time limited
the information concerns relative quantitative parameters (granularity of information on for example files sorted on size)	+	++	+			sound frequency, odour concentration, brightness, pressure, sound intensity
the information concerns absolute quantitative parameters	++	-	- -	- -		
ambient processing (information must be conveyed in the periphery and processed in the background)	- -	+	-	++		
large number of items in sensory or working memory (e.g. extended menus, or menu structure)	+	□				
NOTE: Symbol usage: ++: very well suited to represent the specific information +: good enough, although not optimal □: neutral -: not very well suited, but possible - -: hardly possible or impossible						

5.2.2 Cross modality effects

Multimodal interfaces are not multi-unimodal interfaces. The end result of a multimodal interface may differ from the sum of the components. Three important issues in this respect are:

1) The interaction between modalities. Can a modality affect the percept in an other modality? An example of this is the "blue banana" experiment. In this experiment, people ate apples, pears, and bananas that were all coloured blue and had the same touch structure. The results show that people were not able to recognize the taste, and did not believe afterwards that they ate fruits. This indicates that the end result is not always what you expect.

2) The congruence between modalities. Is information that is objectively the same in two modalities also perceived as the same? Examples of incongruence are for example that an auditive tone seems to last longer than a visual stimulus.

3) The conflict between modalities. When information presented to two modalities differs, how is that perceived? A well known example is visual capture or dominance in the spatial domain. This enables us to fuse the percepts of two modalities in a coherent one based on the visual information. This makes us think that the sound is actually coming from the moving mouths of the actors on the movie screen, and not from the loudspeakers. Also, a similar effect is present in the visual - kinaesthetic relation. For example, if we hold a straight object but see a curved object (e.g. because of glasses, or because the object that is partly above and partly under water such as the stairs in a swimming pool), we identify (and even feel) the object being curved.

Only three aspects of cross modal interactions are studied more extensively: the spatial domain, the temporal domain, and attentional switching. Below we will discuss the different combination of modalities.

Vision and audition

Vision dominates audition in the spatial domain, also known as visual capture (Hay, Pick & Ikeda [118]). Effects are found in for example estimating the magnitude of a length (Teghtsoonian & Teghtsoonian [175]) and in perceived size. However, audition dominates vision in the temporal domain (Handel, S. & L. Buffardi [114]; Geldard [95]; Goodfellow [107]; Lechelt [139], cf. White & Cheatham [192]; Kolers & Brewster [137]). There is an incongruency on perceived duration, i.e. auditive events seem longer than visual (Behar & Bevan [38]; Goldstone & Goldfarb [105]; Walker & Scott [187]). There are no indications that attentional switching between vision and audition involves additional costs (Spence & Driver [168]).

Vision and kinaesthesis

The pattern between vision and kinaesthesis is similar to that of vision and audition. Vision dominates kinaesthesis in the spatial domain, but kinaesthesis dominates vision in the temporal domain (Handel, S. & L. Buffardi [114]). There are also indications that there is a similar incongruency on perceived duration: auditive events seem longer than visual (Werkhoven & Van Erp [191]). Regarding attentional switching: there are additional costs involved for switching attention from touch to vision, but not vice versa. In general, touch and vision already go together in many interfaces. Strong point in this regard is the well-developed eye-hand co-ordination in many users.

Audition and kinaesthesis

In the spatial domain, kinaesthesis dominates audition, but in the temporal domain audition dominates kinaesthesis. Temporal events are perceived congruent across audition and kinaesthesis (Ehrensing & Lhamon [84]; Hawkes, Deardorff, & Ray [117]). Additional costs are involved when switching attention from kinaesthesis to audition. In general, audition and kinaesthesis have more or less the same pros and cons, so the surplus value of the combination may be restricted. However, the combination might be useful to build in redundancy. Please be aware that vibration often will produce sound as a by-product. Adding kinaesthetic information to speech output or auditory icons can be useful because speech information may be very poor or difficult to use, and because kinaesthesis can help to disambiguate the many-to-many mapping of auditory icons.

5.2.3 Lessons learned from the application of multimodal interfaces

In this clause we present some of the results of studies on multimodal interfaces. Although the pool of examples is still small, this is only a limited presentation. A couple of studies improved the mouse and other operations applicable to a windows environment. Brewster, Wright, and Edwards [51] found that an auditory enhanced scrollbar results in faster time to completion and lower mental effort. Akamatsu and Sato [31], [32] found similar results when they added tactile and force information to the mouse which indicated when the cursor was in the right position of the target. Their evaluation experiments showed that additional tactile and force information reduced the response time, made the effective target size large, and reduced the dependence on vision. The advantages of redundant information presentation (in this case cockpit warnings) in choice reaction times is for example provided by Selcon, Taylor, and Shadrake [163]. This advantage is even larger when the original channel is degraded, for example by environmental factors such as a visually cluttered environment (Wogalter, Rashid, Clarke, and Kalsher [195]. However, not all multimodal interfaces result in better performance. Forren and Mitchell [88], for example, found that speech input actually hindered the operator's control performance in a simulated control room environment, although this is a typical work environment where a need exists to develop multimodal interfaces to take advantage of processing resources which are not currently used to advantage. More recent work casts doubt on the effectiveness of speech-based multimodal interfaces for tasks that require extended processing of information and recall of information from memory comes from Cook, Cranner, Finan, Sapeluk, and Milton [65], and Cook, Angus, Campbell and Cranner [64]. These results indicate that the success of a multimodal interface may be related to the task demands. Favourable results of redundant information presentation are found in relatively low level tasks such as mouse control and choice reaction tasks. Adverse effects are present for tasks that require higher levels of cognitive processing. However, no systematic research on this task dependency is currently available.

5.2.4 Guidelines for multimodal symbols

This clause presents the guidelines that can be distilled form the data presented. Please note that the status of the work on multimodal symbols and interfaces is in an emerging phase. That means that most of the guidelines are tentative and have not been thoroughly validated. However, even in the present, preliminary, form they can be useful; both to the designer and to identify interesting areas of future research.

M1. Avoid cross-modal comparisons

Cross-modal comparisons of parameters is less accurate that intra-modal comparisons (Davidson & Mather [78]).

M2. Avoid cross-modal tasks

Cross-modal tasks are more difficult than intra-modal tasks (Conolly and Jones [63]; Millar [145]).

M3. Make the information content congruent across modalities

Modalities must be co-ordinated and made congruent informational.

M4. Make the spatial aspects congruent across modalities

Modalities must be co-ordinated and made congruent spatially.

M5. Make the temporal aspects congruent across modalities

Modalities must be co-ordinated and made congruent temporally, events in different modalities must be well synchronized.

M6. Events must have a natural sequence

Serial events in different modalities must have a natural sequence.

M7. Avoid attentional switching

Multimodality may involve extra costs by cross modal attentional switching.

M8. Allocate information to an appropriate modality

Use a modality to present information it is well suited to present, and allocate each different part of the information conveyed in a symbol to the modality it suits best.

M9. Be aware of conflicts between modalities

Be aware of the potential pitfalls of multimodal symbols: incongruency, sensory conflict, cross modal interactions.

M10. Higher processing levels do not automatically benefit from multimodality

Multimodality may not be beneficial for information that requires high level information processing.

M11. Consistency

Use a consistent allocation of information to a modality, within and preferably between applications.

M12. Benefit

Do not employ multimodality if there is no benefit to the user.

M13. Side effects

Be aware of the side effects of using certain modalities: sound may distract other users, vibration produces sound, etc.

M14. Comparable modalities

Audition and haptics have a lot of strong and weak points in common. Combining them may be easy to provide redundancy.

M15. Complementary modalities

Vision differs substantially from audition and haptics with regard to strong and weak points. This provides possibilities to combine them as complementary information sources.

M16. Benefits for visual symbols

Add multimodality to a visual symbol if it has one of the following attributes:

- timing: add audition or possibly touch;
- 3D localization (e.g. objects behind a window): use audition or touch;
- privacy: add touch;
- outside area of interest: add audition or possibly touch;
- warning/alert: add touch or possibly audition;
- cultural generality: use audition or possible touch;
- ambient processing: add olfaction and possibly audition.

M17. Benefits for auditory symbols

Add multimodality to an auditive symbol if it has one of the following attributes:

- spatial: add vision;
- privacy: add touch;
- physical object: add vision;
- persistence: add vision or touch;
- absolute quantitative parameter: add vision;
- large number of items in echoic memory: add vision.

M18. Benefits for haptic symbols

Add multimodality to a haptic symbol if it has one of the following attributes:

- spatial: add vision;
- abstract: add vision or audition;
- optimized reaction times: add vision;
- memorability: add audition;
- physical object: add vision;
- absolute quantitative parameter: add vision;
- ambient processing: add olfaction or audition.

M19. Alternative channels

The following table gives an overview of the appropriateness of combining modalities as redundant information sources. It can be used as follows: the primary channel is given in the first column (**bold**). The top row lists the possible alternative channels (*Italic*), and the scores indicate the appropriateness. Please note that the table is not symmetrical around the diagonal. For example, if the primary channel is text then the preferred alternative channel is speech, and animations, non speech sound and tactile may possible only convey specific parts of the information.

Table 5.3: An overview of the appropriateness of combining modalities as redundant information sources

alternative primary	te	gr	an	sp	ns	ta	ki	tm	ol	gu
text (te)		-	□	++	□	□	-	- -	- -	- -
graphics (gr)	++		++	+	□	++	++	- -	- -	- -
animations (an)	++	+		+	+	+	+	- -	- -	- -
speech (sp)	++	+	+		□	-	-	- -	- -	- -
non-speech (ns)	+	+	++	+		-	-	- -	- -	- -
tactile (ta)	□	-	-	□	+		+	- -	- -	- -
kinaesthetic (ki)	□	- -	-	□	-	+		- -	- -	- -
temperature (tm)	++	+	+	++	- -	-	-		- -	- -
olfaction (ol)	□	-	□	□	- -	- -	- -	- -		++
gustation (gu)	+	-	□	+	- -	- -	- -	- -	++	

NOTE: Symbol usage:
++: very well suited as alternative output channel
+: good enough, although not optimal
□: neutral
-: not very well suited, but possible
- -: hardly possible or impossible

Annex A (informative): Design for all philosophy and criteria

A.1 What is Design for All?

Our efficiency in interacting with the Information Society depends on the capacities of each individual and on how products and services around us are designed. However, as we grow older, our characteristics and our activities change. When we are children, our size makes it impossible for us to reach, handle or understand a series of objects and products, sometimes for reasons of safety and sometimes because children have simply not been envisaged as potential users. When we have to look after a baby, we find ourselves in all sorts of situations in which we have to cope with using only one hand, we have to use some information services innovative for us and we take good care to alter the way that various features in the home are designed (protecting access to Internet, securing PCs and cupboard doors and so on).

In our lives as adults, we come across countless situations that make it difficult for us to interact with our surroundings (a bandaged cut hand, a leg in plaster, surfing the Internet without our glasses, the discomfort caused by a back-ache or a stiff neck, visiting a country with a different language, etc.). When we reach a certain age, we loose strength and resistance as the years go by, our articulations become less flexible, our senses lose in perception, it becomes more of an effort to remember things and it is more difficult for us to understand innovative products and services.

It is also possible, although less probable, that we will contract some physical, psychological or sensorial disability as our lives advance. This happens to about 10 % of the population under 65 years of age. But if we are lucky and this does not happen to us, then by totalling together all the time that we have problems with our surroundings, because we are children, because we have twisted an ankle skiing or because our abilities are just not the same any more as we grow older, we can calculate that we shall be suffering from difficulties in interacting with our surroundings for at least 40 % of our entire life-cycle.

Why, then, is the way that our environment is designed not better suited to our real needs?

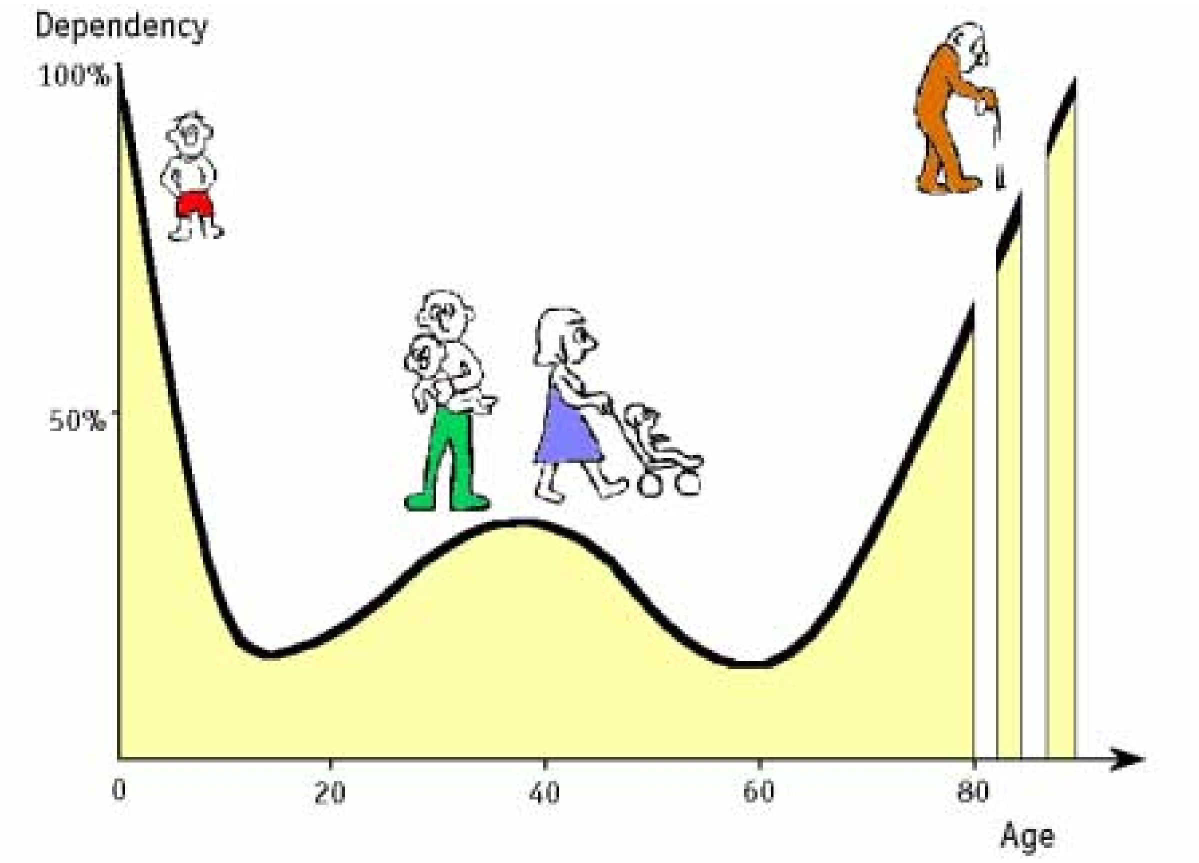

Figure A.1: Dependency on the others as a function of age

While there is no doubt that these questions are seldom taken into consideration during the design process, it is equally certain that we tend to blame ourselves for our awkwardness and our lack of strength. Although we accept the presumption that we are a species that has modified the natural environment in order to survive, we tend to think that it is impossible to make things easier to use and that there is in any case a whole series of inconveniences and misfortunes that we simply have to get used to putting up with.

As we fail time and again to interact with everyday objects, we eventually convince ourselves that we ourselves are "useless", rather than realizing that the objects and services are simply badly designed. Nevertheless, as we live in an environment created by human beings for human beings, we must realize that any problem arising in the way that we interact with that environment is caused by its intrinsic inadequacy for our needs and not by any maladjustment to the environment on the part of our abilities.

One of the reasons for this maladjustment is the overriding tendency to design for a "normal" public or for the statistical mean of the population. We have to remember that what is "normal" for a human being is actually diversity, which is the characteristic that sets us apart as a species, so it follows that it is "normal" that the people who use a product are very diverse, that the uses they make of it will be different from those for which it was originally intended and that misunderstandings will arise about its use if it is ambiguous.

Another mistake that is often made in the design process is the incorrect use of anthropometric tables for the purpose of designing for the statistical mean of the population. However, it may be obvious to keep on repeating it, we have to remember that there is no such person as Mr. Average: he is simply the result of adding up all the different features of all the individuals in a sample and dividing the result by the number of individuals it consists of. Would anybody really even dream that it might feasible to design a vehicle by giving it the exact number of wheels that come out of a calculation of the statistical mean for all the vehicles in any given town or city?

The truth is that we are all more or less distant from any mean parameter, so that, if a product is designed by applying the criterion of the statistical mean, the further away we are from that mean, the more uncomfortable it will be for us to use it. For this reason, it is much more satisfactory to take the dimensions of the individuals to be found at the extremes of the Gauss distribution into account and to try at long last to develop solutions to the dimensional needs of all the individuals to be found between them with just one product, with one that is adjustable or with a range of products.

Another aspect to be borne in mind is that the professional or the entrepreneur seldom contacts users to find out what they really need, what they expect from a product or the evaluation they make of products they already use, so we have the paradoxical situation that, although the countless functions to be found on a video cassette recorder are worse than useless and not at all convenient for practically everybody, they go on being manufactured in such a way that they are inadequate for the majority of users, with such features as:

- displays that are difficult to see from the distance from which television is watched;
- small and badly differentiated keys;
- remote controls with function keys whose purpose we always forget;
- a shape designed to be put under the television, so you have to stoop to adjust it and so on.

As it has been defined, there is always a gap between the products (including the physical environment, products and services) and the human being, because the products do not meet the abilities perfectly. This gap is depicted in figure A.2.

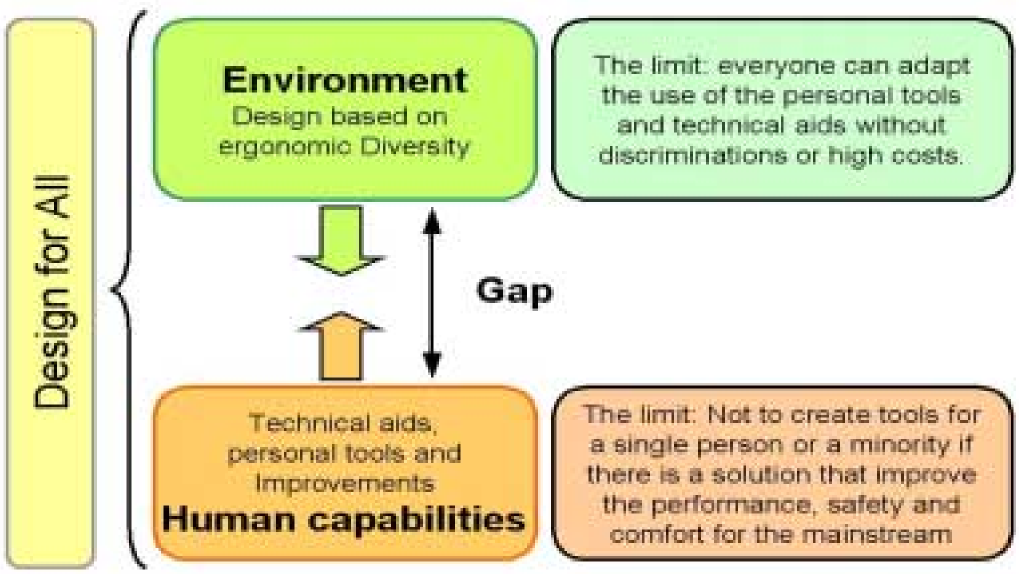

Figure A.2: Gap between products and the human being

In order to modify the environment to adapt increasingly to our needs, thus reducing that 40 % of our lifetime in which we have difficulties, there is a need for a change in the way that products and services are conceived of, so that they will be designed for all.

Design for All means taking the following into account when conceiving products and services:

1) The environment we live in has been and continues to be designed for human beings, by human beings, so it follows that it must be the environment that adapts to our needs and not the other way around.

2) There is an enormous diversity among users' physical, cognitive, sensorial, dimensional and cultural characteristics.

3) Users evolve throughout their lives and their abilities and attitudes change with time. From a state of total dependency when we are babies, we move to one of interdependency with people and our environment that enables us to enjoy distinct degrees of autonomy throughout our lives, depending on our age, individual abilities and capacities, our responsibility for other people, our economic capacities etc. Our dependency on others varies throughout our life cycle. When wee are born, we are totally dependent on our surroundings; by the time we are 40, our parents and children depend on us to varying degrees; then, as we get older, we tend to become increasingly dependent on others.

4) Users are the best source of information for adapting products and services to their needs and expectations.

So, we could define Design for All as the interventions on environments, products and services with the aim that everybody, included the future generations, independently of the age, sex, capacities or cultural baggage can enjoy participating in the building of our society.

It is a simple idea: make products and services usable by everybody. It serves two purposes at the same time: meeting the needs of consumers who have difficulty using some products, and meeting the needs of companies who want to expand their potential market. Due to the flexibility of the new ICT products and services, the industries of this sector will discover new methods, devices and services that will increase the participation possibilities to many people excluded until now.

A.2 Design for All legislation and standards

A.2.1 A comparison between USA and European approaches

In general, the European approach to legislate on Design for All matters has been based on promoting consensus between countries, on funding research initiatives and issuing mandates to relevant technical and organizational bodies. On the contrary, the USA has issued concrete legislation in several areas, through the "American with Disabilities Act" and the Telecommunications Act of 1996, and Governmental Agencies which enforce these laws. A more detailed review of both sets of laws and directives can be found in Appendix 2 of the present document.

In summary, the basic differences are:

- The EC has promoted consensus and issued general directives, while the US have made concrete laws addressing the problem. This effectively has created a broad mandate for accessibility in the US, while in Europe there is still the need to harmonize legislation and standards.
- Rather than addressing inaccessible technology already in place, US law addresses technologies while they are still under development. For example, the Telecommunications Reform Act (1996) requires all telecommunication manufacturers and service providers to ensure that new products are accessible or compatible with assistive technology where readily achievable.

The EC recognizes the effectiveness of the US approach in the "Mandate to the European Standards Bodies for Standardization in the field of information and communications technologies (ICT) for disabled and elderly people" Brussels, 6 May 1998 DG XIII-C5/DR/D(97)" with the following words:

- "A likely effect of this legislation will be that US industry will be better placed to respond to the demographic changes in society and corresponding changes in the market".

A.2.2 Design for All and HF Standards

There are two main difficulties about the application of the Design for All approach in the ICT industry: the scope is enormous and has no clear boundaries, and, in addition, no pan -ICT standards currently exist with regard to the Design for All approach [18].

The speed at which ICT is developing and the convergence between technologies will require organizations to work more effectively across traditional oganizational boundaries if he recommendations are to be implemented, and a digital divide avoided. Appropriate standards will be required so that the aim of Design for All can be achieved.

The following clauses in this part review the current state of these standards.

A.2.2.1 The draft ISO/IEC Guide 71 on the needs of older persons and people with disabilities

The ISO/IEC GUIDE 71 [16] provides guidance to writers of International Standards on products, services and environments as to how to take into account the needs of older persons and people withdisabilities. The purpose of ISO/IEC guide is:

a) to raise awareness about how human abilities impact on the usability of products, services and environments,

b) to outline the relationship between the requirements in standards and the accessibility and usability of products and services, and

c) to raise awareness about the benefits of adopting universal and accessible design principles.

The main part of this ISO Guide provides a review of human abilities and limitations with the aim that they are used for the preparation of other standards.

Regarding the role of Design for All, this guide states that standardization greatly influences the design of products and services that are of interest to the consumer and therefore can play an important role in this field. However, this needs to be considered within the constraint that standards should normally not be design-restrictive.

The ISO guide also recognizes that general guide cannot provide comprehensive information for specific product or service sectors and additional sector-related guides may need to be developed.

A.2.2.2 ISO Standard 13407 on Human-Centred Design Process

The most relevant standard about Human-Centred Design Process is ISO 13407 [197], of 1999, on the Human-centred design process for interactive systems [27]. The present document does not deal with individual user interface elements or techniques, but rather, with the process involved designing systems with the human in mind. The standard specifies activities necessary to ensure human-centred design of interactive computer-based systems - both hardware and software. It provides a framework for reporting that all necessary activities have taken place. It is a tool for those managing the design processes and provides guidance on sources of information and standards relevant to the human-centred approach. The standard covers topics such as:

- Rationale for adopting a human-centred design process;
- Principles of human-centred design;
- Planning the human-centred design process;
- Human-centred design activities.

The standard was published in 1999. The present document was fully referenced and described in EG 201 472 [29].

A.3 Basic principles for achieving Design for All

Just as we proposed in the previous parts, if we want a product to suit its users' abilities and needs, we have to proceed through the product or service development process, taking a whole series of criteria into consideration that will guide us throughout the relationship that the product or service will have with its users.

These criteria are as follows:

- Facilitating use of the product. By product we mean environment, products and services.
- Ensuring that the users take part in the product design and evaluation process, so that, for a representative sample of potential users, the product or service will suit their anthropometric and functional characteristics and will in its turn be compatible with their habits and culture.

According to these criteria, the designer must adapt his working methods so that they will take the users into account in each phase of the design process.

It is worth remembering that, for the majority of products and services that are launched onto the market, the user may:

- be a man, a woman or a child of any age;
- have an asymmetric body;
- have reduced sight or be blind;
- have reduced ability to hear the spoken word;
- have a restricted cognitive ability or understanding of language;
- have a restricted memory capacity;
- have slow reflexes;
- have difficulty grasping things in his/her hands;
- have only one hand or no hands at all;
- use prosthesis, an orthesis or a technical aid;
- walk with crutches, walking sticks, a walking frame or other equipment;
- have difficulty crouching down or standing up;
- be allergic to various materials or to static electricity;
- be technological illiterate or non western-culture.

For this reason, the designer must make the effort to banish the idea of imagining himself to be the model of all possible users. The designer must have into account that the use of the product or service being designed necessarily call for a series of previous and subsequent steps, in other words, before using it, users will have to know that it exists, acquire information about it, make the decision to use it and manage to reach it.

All this necessarily presupposes that the designer provides solutions so that all these steps previous and subsequent to using the product or service are suited to its users' characteristics and needs.

As a conclusion, all designers should be aware of the following principles for achieving Design for All.

A.3.1 Facilitating use of the product

Facilitating the use of the product means taking seven premises into account:

1) **Simplicity:** superfluous elements and operations must be reduced to a minimum.

2) **Flexibility:** the design must adapt to the users' abilities to interact with it. Its use will therefore have to be flexible enough to adapt to its users' characteristics.

3) **Quick information:** the system must enable users to perceive quickly and unequivocally what it is and how they should start using it.

4) It must respond to a **conceptual model** of functioning that adapts to users' previous experience and expectations.

5) There must be a **clear relationship** between the activation systems at users' disposal and the results that they generate.

6) It must contain a **feedback** system that keeps users permanently informed about the product condition and what it is doing.

7) **Error prevention and handling:** users may misunderstand or use the product for a purpose other than the one for which it is intended, without this causing any harmful consequences. The system must provide mechanisms for the user to solve this situation.

A.3.2 User participation in the design process

The starting point of the concept of Design for All is the principle that all potential users can use the product or service correctly and that people find the product easy to use. User participation in the design process provides direct information about how people use products.

Nevertheless, direct contact between users and designers can be risky if it is not properly structured. Designers can use interviews to strengthen their own ideas or collect piles of unstructured data. User participation should be structured and systematic, starting by formulating the specific aims of involving the users in the design process.

The information provided by the user participation is often very profitable and necessary for constructing objective arguments for the decisions that are made during the design process, although the evidence may be costly. A full review of methods for the involvement of users in the different design phases is presented in EG 201 472 [29].

A.4 When to apply different Design for All focusing

To sum up, we can conclude that there are three main approaches of how to apply the Design for All criteria to the design of products and services and, therefore, to the design of multimodal symbols.

- **A solution for all:** when an open-to-public system that we are designing should be used sporadically by everybody, we should take into account those requirements including human diversity. So, everybody must use all those services that we can find in different public spaces, streets, buildings or transport, without an expensive adaptation.

- **An adaptable solution:** The system that is being designed must fit the user requirements when is a public service but with a long operation time or services operated in private environments. That is to say, in all those public services where the request has a long operation time, as for example a query of transport timetables, or a service that can be requested from a private environment as the office or home. In these cases the system should provide the user with the possibility to configure the system to be adapted to his/her specific abilities: modify the size of symbols and letters, incorporation of sound, modify contrast.

- **A range of solutions:** it must be guaranteed that the market has a range of solutions that allows the selection of a model usable by an specific user disregarding their capabilities, when existing technology does not allow to use one of the above mentioned approaches. In this case, we should wait technological improvements so those mainstream products could be designed. We could find this solution in those services that user can find in public spaces but mainly it is a solution for those services that are used in a more private way.

Table A.1shows a summary of the three approaches with some application examples:

Table A.1: Summary of recommendations about when to apply different Design for All focusing

CRITERIA	WHEN TO APPLY	EXAMPLES
A solution for all	Public services with a short operation that should not need any adaptation	Signalling and guidance system in the thoroughfare and roads. Signalling system in a transport mean: airport, underground or train station. Informative screens used in public spaces
An adaptable solution	Public services with a long operation time or services operated in private environments	Public access terminal should be adaptable to user's abilities, with systems for blind navigation or language change. Websites can have an special link to a text version. Graphic operation systems for personal computers should provide user with a configurable environment.
A range of solutions	When current technology does not allow the use of one of the previous criteria	ICT personal devices that empower different characteristics like autonomy, size, weight instead of friendliness. Examples are PDA or mobile phones.

Annex B (informative): User requirements for multimodal symbols

B.1 Users and their diversity

To implement a DfA approach to the design of ICT products and services, and multimodal icons, symbols and pictograms in particular, it is necessary to understand the diversity of users, their abilities and capabilities, and the limitations they may have in their abilities and capabilities, whether they be temporary, long-term or permanent. The purpose of this clause is, then, to outline the various sensory, physical and cognitive abilities and capabilities of human beings, their possible limitations and the effects of ageing, as a basis to provide user requirements for multimodal icons, symbols and pictograms, taking into account the greatest possible range of users.

There are a large number of attributes that can be used to distinguish between people in a population. The ones that should be considered to have direct impact on the successful use of ICT products and services [197] include:

- **Sensory abilities** such as seeing, hearing, touch, taste, smell and balance.
- **Physical abilities** such as speech, dexterity, manipulation, mobility, strength and endurance.
- **Cognitive abilities** such as intellect, memory, language and literacy.

The individual user may have excellent ability in some areas and yet be poor in others. For the population as a whole there can be a wide variability in any one attribute. The complexity of the problem increases dramatically as more attributes are considered.

For the purposes of this annex, only the relevant human abilities and disabilities will be reviewed. The interested reader can consult EG 202 116 (see Bibliography) for a more detailed review of them for the widest range of ICT product and services

The basic human abilities required for an effective use of modern ICT products and services are:

- Sight;
- Hearing;
- Handling; and
- Cognition.

B.2 Sight

Sight is essential for being able of using modern ICTs, being most of them, if not all, based on visual displays of information. Sight is one of the intrinsic human abilities that differentiate us as a species, and at the same time there is a huge variety of potential disabilities, caused by many reasons, and affecting any person at any age. Estimates exist that a 3 % of the whole population suffer any kind of visual disability. This of course depends on the exact definition of "visual disability". For instance, for being recognized legally as a blind person, the vision ability should be less than 10 %.

It is precisely because of the necessity of this ability and the complexity of the disabilities that designers should be aware of the whole range of them, so that they can provide mechanisms for the different disabled users to customize, or to adapt the products to their specific needs.

Users with any sort of visual disability are among the most benefited users of the introduction of any multimodal capability in user interfaces. Although it is clearly an over-simplification of a very complex issue, the disabilities will be classified as a function of its effect in the use of ICT devices in 3 levels: blindness and tunnel vision, disabilities in one property of sight (e.g. color, movement or contrast), and optical disabilities. A thorough review of the different alterations of sight can be found in Internet at http://www.rnib.org.uk/info/welcome.htm.

B.2.1 Blindness and tunnel vision

The most severe disability is blindness or the loss of central vision. There are different degrees of blindness, which are classified in terms of the perception of light: some people do not perceive light at all, others can distinguish between light and dark, some can perceive slight movements or some images. Loss of sight can involve one eye (leaving monocular vision) or both eyes. A person is generally only considered to be "legally blind" when his or her sight is reduced to 10 % of perfect vision.

Any form of blindness of loss of central vision makes such activities as reading or writing very difficult, if not impossible. Most usually, blind people cannot effectively use ICT products displaying information in visual displays, as PCs, PDAs, or an agenda in a mobile telephone, for instance. These users require then that special assistive products can be used together with the product, as Braille input and output devices, or providing alternative ways of rendering and introducing the information (e.g. speech synthesizers and recognizers). Braille is an excellent mechanism of presenting information, but only for those trained in using it (estimated at around 10 % of the population of blind).

Other severe sight disability is tunnel vision or partial loss of peripheral vision. Tunnel vision consists of a dramatic reduction in the vertical and horizontal angle of the peripheral field of vision; the effect is practically as though you were to be looking through a hollow tube or a tunnel. The only objects you can see are those that are situated immediately in front of you. This condition usually arises from the age of 35 onwards, although a more frequent form is found after 75 years of age.

In the partial loss of peripheral vision, the field of sight is reduced, narrowed down, so you cannot perceive objects situated at the extreme of your field. Depending on the degree of the loss of peripheral vision, these disabilities do not, in principle, make impossible to use visual displays of information. However, the modern trend to present large quantities of information in visual displays presents many difficulties for them. Designers should be aware of this disability, and provide mechanisms to customize the interface for these users, for instance, screen magnifiers, large size in text fonts, and the possibility to manipulate the presentation of the information so that it falls within the field of view of these persons (e.g. zoom capabilities) and can scroll in all directions of complex screens (e.g. virtual desktops). Also, and depending on the particular disability affecting the user, maximizing contrast greatly benefits these users. The type of contrast should be configurable by the user, and as a minimum two modalities should be provided: black information over white background, and white information over black background.

People with these disabilities are the ones most benefited from the introduction of multimodal capabilities in the user interfaces.

B.2.2 Color disabilities

Approximately 7 % of men and 0,05 % of women have some form of colour blindness. Unfortunately, there are five different types of colour blindness with different consequences for the colour discriminations which are possible. The most common form is deuteranopia, also known as red-green deficiency, in which light in the "green" area of the wavelength is responded to, but green cannot be distinguished from certain combinations of red and blue. Therefore colour contrasts between violet and blue or within the green/yellow/orange/red segment of the colour spectrum should be avoided.

Perception of colour is also affected by the general ageing process and a number of diseases and disorders such a diabetes and glaucoma. In all these cases, the loss of colour perception is towards the violet/blue end of the colour spectrum. Therefore fine contrasts at this end of the colour spectrum should be avoided.

The general recommendation for meeting all these users' requirements is to avoid using colour as the only or the essential property of the functions to be distinguished, ensure that the information can be used in monochrome. Some of these users may also require functions to maximize the contrast in the user interface, as recommended in clause B.2.1.

B.2.3 Optical alterations in the eye and other vision disabilities

Other disabilities affecting sight are cataracts (an opacity of the crystalline lens), halo vision (the perception of a halo of colours around a light source and is one of the symptoms that indicates the presence of a glaucoma, a cataract or conjunctivitis), blurred sight (the loss or deformation of sight or the appearance of confused, blurred or hazy images), and flies (mobile shadows that are perceived in the visual field and make it difficult to see objects or read). For all these disabilities, the recommendation is to provide mechanisms to maximize the size what is presented in the display, and alternative ways of presenting the information.

There are also very frequent alterations in sight due to optical alterations in the eye. Their most common effects are difficulty in seeing objects at a distance, in seeing nearby objects, and astigmatism. Users with these disabilities, when properly corrected, and unless very severe, should be able to use an ICT product normally. However, it is important to provide for mechanisms to present information in larger sizes so that these users are more comfortable for viewing it at a normal distance.

Other frequent problems is the difficulty in adapting to light. This occurs when we pass from a dark area to one that is well lit or vice-versa, we are unable to see properly for a few moments, until the cones and rods have adapted to the new situation. The effect of certain infirmities and of ageing is that the time it takes the cones and rods to adapt gets much longer, so that proper sight sometimes does not return for several minutes. This should be taken into account in applications in which it is foreseen that such a change in light conditions will happen. For instance, in cash-dispensers that are located inside buildings in a sunny day. In that case, provide for additional lighting to minimize the effect.

A related problem is experimenting difficulties in seeing details in conditions of poor lighting. The most frequent reason for this phenomenon, is one of ageing: as the body ages, the lenses become more opaque and light and colours are perceived as though they have passed through a yellow filter.

Apart from these disabilities, we should also consider temporary alterations, a whole series of factors that can make us lose our visual ability at any time, including:

- The use of sunglasses.
- Dazzling glare produced by a mirrored surface or a fire.
- Losing your glasses or dropping a contact lens.
- Having an eye covered over for medical treatment.
- Artificial light suddenly going off at night.
- Lights in a discothèque, party room etc.
- Your glasses getting steamed over when you are cooking or raindrops accumulating on the lenses.

B.3 Hearing

Hearing is another basic human ability, and essential for human communication, as it is the basis for speech communication. Although the different forms of hearing disabilities do not prevent the affected people from using ICT devices, they clearly affect their capability to communicate with others, and the most obvious example is using the telephone service.

Hearing disabilities have different degrees of severity, ranging from total deafness, other losses (people known as "hard of hearing"), losses produced by the decrease of the ability in certain domains (e.g. the loss of sensibility to specific frequency range due to ageing), and a whole range of temporary losses. Estimates about the percentage of population suffering of hearing disability are around 8 %.

The human voice produces sounds whose frequency is from 2 000 Hz to 3 500 Hz, and it is the loss of perception of these frequencies the one producing the inability to understand spoken communications. People suffering deafness and those with losses in this range of frequencies require for using ICT products that the information is also presented in text or other visual format, and for control information, alternatively using haptic or vibratory information.

Many deaf people use sign language for their communication. This language presents many problems for communication as it is generally not standardized, there are large international variations, and most non-deaf people cannot speak or understand it. For telecommunications, these people require video communication (text communication almost completely eliminates all the emotional information in language), or special relay-services, in which an operator translates the communication from signs to voice, and vice-versa.

Deaf and hard of hearing people greatly benefit of lip reading, up to the extent that some of them can understand a conversation by looking at the speaker when they cannot perceive any sound. A very fruitful line of research is being undertaken to synthesize look-alike human faces that simulate the lip and facial movements of a normal speaker, also known as "avatars", so that sign-language speakers, and also hard of hearing people, can benefit of all the complementary information of face-to-face communication without the bandwidth and terminal requirements of real-time video communication.

For the less severe forms of hearing difficulties, and also when suffering temporary hearing difficulties, designers should consider the specific frequency loss or environmental condition for using the ICT product, and incorporate in the user interface capabilities to customize and adapt the sound properties of the system to the new requirements (e.g. volume control, configurable alarm signals, configurable tones for messages, etc.).

Multimodal formats of presenting information are of great benefit in all cases. Most obvious (and most usual for all of us) is not hearing our mobile telephone ringing in noisy environments. Vibration capabilities usually found in most modern terminals allow us to sense it. This 1-bit information, however, greatly wastes the capabilities of the multimodal information. Guidelines presented in clause 7 are intended to enrich the uses of these capabilities now present or possible in many ICT products, as, for instance, computer mouse, trackballs, joysticks, even keyboards.

B.4 Handling and co-ordination

Other basic human ability required for effectively using any ICT product or service is manual handling. All devices require some form of manual operation, as pressing keys, touching parts of a screen, turning knobs, etc. This is a human ability greatly affected by ageing. Ageing frequently has its effects on the brain or brings on infirmities in the central nervous system such as Parkinson's Disease, while trembling, which is often found among the elderly, makes all these everyday gestures more difficult to undertake.

A shortlist of disabilities that can be found under this category are:

- Difficulty with the pincer grip;
- Difficulty with gripping;
- Difficulty in controlling small movements;
- Limited articular movements in the hands;
- Temporary alterations in handling ability.

All these problems affect the use of input devices of any ICT device, not so the output devices.

B.5 Cognitive abilities

All the above capabilities obviously require cognition. Cognition is a process of mental representation whereby human beings perceive, imagine, classify, conceptualize, formulate and solve problems in the framework of their personal experience and cultural environment.

Cognitive problems are found among the elderly, people with psychic disabilities and people who have had a vascular cerebral disorder of varying significance, bringing alterations that affect the individual's social and professional behaviour. In addition, certain psychological states, such as depression, stress or other psychic alterations, can change our ability to perceive and process information quite radically.

A list of potential cognitive disabilities is the following:

- Alzheimer's disease;
- Difficulty with orientation in space and time;
- Difficulty with memory processes;
- Mental disability;

- Difficulty in remembering all the steps in complex operations;
- Difficulty in understanding complex processes;
- Inability or difficulty with speech;
- Temporary alterations in cognitive abilities.

Apart from these more or less physiological disabilities, we have to take into account culture, customs and habits as something having a great effect on the use of ICT.

Customs and habits are forms of behaviour or ways of doing things that we acquire by learning, especially by repetition, so much so that they almost become automatic. They are acquired forms of behaviour typical of any social group, regardless of its extent, or of an individual.

Users can experiment difficulties when they do not have enough comprehension or experience with specific cultural habits, for instance, when in different socio-cultural contexts, when having to use a different language, etc.

Multimodal information, as an alternative and redundant way of presenting information, can greatly help in all this situations. A basic property to take into account is to make the multimodal information not dependent on specific language, cultural or contextual setting, and trying to make it understandable by as many different people as possible.

Annex C (informative): Relevant guidelines and standards about symbol usage in current ICT applications and services

In order to produce guidelines for future multimodal symbols, it is essential to have a good understanding of how symbols work in current ICT applications and services. This state of the art report has then the following purposes:

- to review current standards for icons, symbols and pictograms;
- to identify the most common types of symbols in current ICT applications and services;
- to review requirements for the accessibility of symbols in current ICT applications and services. In particular, for Web services, the Web Accessibility Guidelines (WAI).

C.1 Current Standards on icon and symbols

C.1.1 ISO/IEC Standards on icons and symbols

ISO and IEC have produced and are still working on the ISO/IEC 11581 standard on "Information technology - User system interfaces and symbols - Icon symbols and functions". This international standard comprises six parts. These parts apply to icons that are displayed on computer screens. These icons represent data objects or computer system functions that users can manipulate and interact with. The parts are:

- **Part 1: Icons - general**

  Provides a framework for the development and design of icons and their application on screens. This part contains general requirements and recommendations for icons and global variations to the graphical representations of icons.

- **Part 2: Object icons**

  Describes user interaction with and the appearance of icons that represent functions by association with an object and that can be moved and opened. This part contains requirements and recommendations for 19 commonly used object icons.

- **Part 3: Pointer icons**

  Describes user interaction with and the appearance of icons that are logically attached to a physical input device, and that the user manipulates to interact with other screen elements. This part describes user interaction with and appearance of pointer icons on the screen. It also specifies how pointer icons on a screen change appearance to give users feedback. This part contains requirements and recommendations for 8 commonly used pointer icons.

- **Part 4: Control icons**

  Defines user interaction with and appearance of the graphical elements that provide task control for the user of the computer display. These control icons can be used to operate on windows, lists, and other graphical elements that provide dialogue interaction between the system and the user. This part contains requirements and recommendations for 14 commonly used control icons.

- **Part 5: Tool icons**

  Describes user interaction with and appearance of tool icons on the screen. It also specifies the relationships between tool and pointer icons. Part 5 contains requirements and recommendations for 20 commonly used tool icons.

- **Part 6: Action icons**

Describes user interaction with and the appearance of tool bar or "action" icons. Action icons represent actions by association with objects that prompt the user to recall the intended actions. This part contains requirements and recommendations for 23 commonly used action icons.

Table C.1 lists briefly the icons that are included in ISO/IEC 11581-2, ISO/IEC 11581-3 and ISO/IEC 11581-6.

Table C.1: Icons described in ISO/IEC 11581 (parts 2, 3 and 6)

ISO/IEC 11581-2		ISO/IEC 11581-3	ISO/IEC 11581-6	
Document	Calendar	Default pointer	Help	Increase indent
Folder	Calculator	Text pointer	Find	Decrease indent
Filing cabinet	Clock	Border control pointer	Create new	Enumerate/numbering
Mail	Display	Cross-hair pointer	Open	Itemize
Printer	Keyboard	Busy indicator	Save	Make bold
Telephone	Mouse		Print preview	Italicise
Facsimile	Network		Print	Underline
Diskette	Audio device		Cut	Align left
Waste paper			Copy	Align right
Can			Paste	Centralize
			Undo	Full justify
			Redo	

ISO/IEC 11581 standard focuses on information systems in general, and some of these symbols are quite specific to document processing systems. Communication technology is not as well represented, and a number of symbols used in current communication applications are not present. Examples of such symbols are "Back", "Forward", "Go to URL", "Go to start page", "Refresh" and "Stop", i.e. symbols and functions that one will find in most web browsers. Symbols specific to e-mail are also not included, e.g. "Send", "Reply" and "Forward".

The operating systems on modern computers (e.g. MS-Windows®, MacOS®) use combinations of graphical symbols, animation and auditory information to represent system objects and actions (e.g. sound of crumpling paper to indicate document deletion, animation of document crumpling and going into waste basket, change of waste basket symbol to indicate that something is in it). These systems are most often configurable; so that the user can set to which extent the different modalities should be used.

The operating systems are also strongly guided by design for all requirements, which means that they should provide means for all types of users to customize the system to adapt it to their specific needs.

C.1.2 ETSI Guidelines and standards about Design for All and symbols, icons and pictograms

ETSI Technical Committee Human Factors has long investigated on the use of icons, symbols and pictograms in several areas of telecommunications. On the other hand, this Technical Committee has also been responsible of providing guidelines about the Design for All approach in this particular field. Some of the past TC-HF documents in both areas are:

- In 1991, ETR 029 [1] on Access to telecommunications for people with special needs was published. ETR 029 provides a review of the different disabilities and guidelines to cope with them for the terminal design and the service operation. The work was updated and particularized for telephone keypads and keyboards in ETR 345 [7], which presented the requirements for elderly and disabled people.

- In 1993, several studies initiated within TC-HF produced as result several documents about pictogram design and evaluation. These are ETS 300 375 [2], which proposes standard pictograms for point to point videotelephony, ETR 070 [3], on the Multiple Index Approach for the evaluation of pictograms, ETR 113 [4] which presents the results of an evaluation study of pictograms for point to point videotelephony. This work was also approved by ITU, in its document F.910 [13]. ITU has also published E.121 [12] with pictograms, symbols and icons to assist users of the telephone service.

- TC-HF work on symbols, icons and pictograms was updated in 1998 with the publication of EG 201 379 [10], "Framework for the development, evaluation and selection of graphical symbols".

- TC-HF has also dealt with activities regarding touch, for instance, by providing a recommendation for a tactile identifier on machine readable cards for telecommunications terminals (see ETR 165 [5] and ETS 300 767 [8]), and also evaluated and recommended a set of symbols to identify telecommunications facilities for the deaf and hard of hearing people.

All these documents and references have something in common: they deal with elements that communicate or are perceived in a single perceptual channel: symbols and icons by means of vision, the tactile identifier by means of touch, etc. The present document does not deal with unimodal navigation elements, but with multimodal elements, i.e. those that are presented and communicate with the user in different perceptual channels or modalities.

C.2 Symbol usage in current ICT services and applications

Icons, symbols and pictograms are visual indication elements which are used in many systems, substituting written explanations, indications, labels or textual commands. These visual representations vary from photographic realistic pictures to representational sketches, line drawings or abstract designs.

Icons are indication elements in the user interface for computer systems or in telecommunication devices. For the former, they are characteristic of Graphical User Interfaces (GUIs), and have special features for its use with these systems. In the latter, they have been traditionally used as an alternative or additional indication to written labels for keys, buttons or other system elements, and also as graphical elements in the visual displays with which most of modern telecommunication terminals are equipped.

The types of applications and services that we will be concerned with in this review are:

- Information retrieval (e.g. Web sites).
- Messaging and text conferencing (e.g. email, SMS, IRC, chat, instant messaging).
- Real time communication services (e.g. fixed and mobile telephony).

General purpose IT applications (e.g. document processing, image processing, accounting applications) are considered outside the scope of the present document. However, a number of the symbols and conventions for use of symbols in general IT applications will also be relevant for the types of applications and services that we review because many of the applications share the interface style of particular operating systems.

C.2.1 Information retrieval (Web) services

When we talk about information retrieval service these days, the name is the World Wide Web, or in short, web. For this basic service there exists an excellent set of accessibility guidelines, that should be known by all web service designers and providers: the Web Accessibility Initiative (or WAI) guidelines [196].

The WAI is an initiative promoted by the World Wide Web Consortium's (W3C), and, among other technical activities, has produced the state-of-the-art guidelines for accessibility of web services. Other excellent guidelines are available in their site (http://www.w3.org/WAI/Technical/Activity.html), including guidelines for user agents. The latest version of the accessibility guidelines is version 2.0, released August 2001.

C.2.1.1 The Web Accessibility Initiative (WAI) Guidelines

For people with disabilities, the accessibility and usability of the WWW is dependent on several components of the Web technology. The structure of information available on the WWW must allow the information to be presented in more than one form and to have it organized in a way that supports the underlying meaning and structure of the information:

- Text information has the capability to be viewed as text, speech or Braille.
- Pictures and movies need text description annotations. A blind person using speech output would benefit from knowing if the current text is labelled as a header or a paragraph.

- The WWW browser need to have features that support the navigation and presentation of information by people with a variety of movement, sensory and cognitive capabilities.

HTML is the most frequently used format for Web documents. It is possible today to create a HTML document that is accessible by almost anyone. The solution is simple: use only text and hypertext links with ASCII characters, avoid graphics and sounds. Then the information can be navigated and read by a text-based browser such as Lynx and a character-based assistive device, such as a Braille display or a speech synthesizer.

But the Web designer does not need to be restricted to text only. There are a number of strategies that can be used to allow the use of graphics and sounds and still maintain accessibility.

This includes images, graphical representations of text (including symbols), image map regions, animations (e.g. animated GIFs), applets and programmatic objects, ASCII art, frames, scripts, images used as list bullets, spacers, graphical buttons, sounds (played with or without user interaction), stand-alone audio files, audio tracks of video, and video.

The WAI guidelines, in its version 2.0, are organized in 4 main guidelines, each of them with several checkpoints. At the same time, these checkpoints are prioritized in levels, providing a basis for conformance testing in 3 levels, described in the next part of this Annex.

The 4 main guidelines are:

- Guideline 1 - Presentation. Design content that allows presentation according to the user's needs and preferences.
- Guideline 2 - Interaction. Design content that allows interaction according to the user's needs and preferences.
- Guideline 3 - Comprehension. Make it as easy as possible to use and understand.
- Guideline 4 - Technology considerations. Design for compatibility and interoperability.

The designer should refer to the latest version of these guidelines, available at http://www.w3.org/TR/WCAG20/. This is a summary of the recommendations for web page designers:

- **Alternative text description for images (ALT tags)**

  HTML provides the ALT attribute (use of ALT-TEXT, text anchor or an alternate text-only page) on IMG elements so that if images are used to control the exact appearance of buttons, logos, etc, there is still standard text available so text-to-speech technology can be used as in screen readers that blind and visually impaired computer users employ. Otherwise the words embedded in the pixels of the GIF or JPEG image are not recoverable as text to be spoken by the screen reader.

- **Updated text-only pages:**

  Text-only pages must be updated in parallel to the main graphics pages. A fully accessible text is of no use, if the information content is obsolete.

- **Sound files:**

  Make a link to a page with a transcript or a description of the sound file.

- **Movies:**

  For movies, make a description of the sound and words of the movie. Use captions, text tracks, an alternate text file or an alternate sound track.

- **Image maps:**

  Image maps allow a user to click on different parts of a picture to reference different WWW pages. This feature requires the ability to see, is completely inaccessible to blind persons. They do not know what the picture is, and do not know where to click, even if the picture is described. Thus, provide text anchors for all links accessible through an image map or, preferably, provide an alternate text-only page.

  Users should be able to switch back and forth easily between text-only and graphic versions of the page.

- **Forms:**

  Forms are usually inaccessible. Provide a form which can be downloaded then mailed or e-mailed, or a phone number where the requested information can be provided.

- **Tables:**

  Minimize the use of tables. Tables cause problems for screen access systems, since the screen reading software tend to read line by line, character by character, from left to right. In a multicolumn document, the screen reader has no logical construct to follow, which makes it very difficult for the user.

- **Use of standard HTML formats, tags:**

- Use standard HTML formats, tags etc. Assistive devices must assume that standard features are used.

- The PDF (Portable Document Format) is increasingly popular on the WWW because of its appealing visual appearance. PDF is largely inaccessible to visually impaired people, because a PDF source document provides no internal element descriptors that can be easily accessed and subsequently translated for a blind user. However, the vendor, Adobe Inc., is attempting to make PDF more accessible. Adobe has developed a special "plug-in" that presents an alternative view of an open document in a separate window. This view contains the texts in as close to reading order as possible.

- Avoid non-standard data structures and viewers. The only way for custom data and views to be accessible is if the access is built directly into the viewer. Standard access tools do not generally work with special viewers.

- Use colours and background patterns that contrast well with the text. Use colours that will makes the pages easy to read by people with colour blindness.

- Minimize the number of multiple hypertext links that appear in a single line of text.

Assistive technology

Assistive technologies for computers are also called accessibility aids. These aids are added to computers by people who use them to make computers more accessible. Some common aids include:

- **Screen enlargers**

  These help people with low vision. Also called screen magnifiers or large print programs, these utilities are like a magnifying glass. People using them are able to control what area of the computer screen they want enlarged, and can move that focus to view different areas of the screen.

- **Screen reviewers**

  They are for people who are blind. These aids make on-screen information available as synthesized speech or a refreshable Braille display. Also called blind access utilities or screen readers, they can only translate information that is text. Graphics can be translated if there is alternative text describing the visual images.

- **Voice input aids**

  These assist people with mobility impairments. Also called speech recognition programs, these enable people to control computers with their voice instead of a mouse or keyboard.

  On-screen keyboards are used by people who are unable to use a standard keyboard. An on-screen keyboard lets people select keys using a pointing method such as pointing devices, switches, or Morse-code input systems.

  Keyboard filters are used by people who have trouble typing, or who want to increase typing speed. The filters built into the Windows and Windows NT operating systems compensate somewhat for erratic motion, tremors, slow response time, and similar conditions. Other types of keyboard filters include typing aids, such as word prediction utilities and add-on spell checkers.

Alternative input devices allow individuals to control their computers through means other than a standard keyboard or pointing device.

C.2.1.2 Levels of achievement of WAI guidelines

Each checkpoint has a priority level assigned by the Working Group based on the checkpoint's impact on accessibility.

Priority 1: A Web content developer **must** satisfy this checkpoint. Otherwise, one or more groups will find it impossible to access information in the document. Satisfying this checkpoint is a basic requirement for some groups to be able to use Web documents. If all Priority 1 checkpoints are satisfied the level of conformance is defined as **Level "A"**.

Priority 2: A Web content developer **should** satisfy this checkpoint. Otherwise, one or more groups will find it difficult to access information in the document. Satisfying this checkpoint will remove significant barriers to accessing Web documents. If all Priority 1 and 2 checkpoints are satisfied the level of conformance is defined as **Level "Double-A".**

Priority 3: A Web content developer **may** address this checkpoint. Otherwise, one or more groups will find it somewhat difficult to access information in the document. Satisfying this checkpoint will improve access to Web documents. If all Priority 1, 2 and 3 checkpoints are satisfied the level of conformance is defined as **Level "Triple-A".**

The WAI web site provides icons for web sites to show their conformance to the different levels provided in the WAI guidelines, as shown in figure C.1

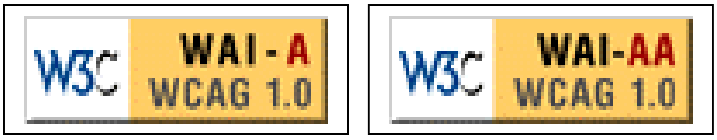

Figure C.1: Graphics used to define "Conformance Level A" and "Conformance Level Double-A"

When designing and developing web pages, accessibility should be validated and ensured. To do it, we can use automatic tools and also human review. **Automated methods are generally rapid and convenient but cannot identify all accessibility issues. Human review can help ensure clarity of language and ease of navigation. An automated tool is reviewed in clause D.2 of the present document.**

C.2.2 Messaging and text conferencing

Instant messaging (IM) has now gained a significant audience, and is interesting from a symbol point of view. Even though the actual messaging is textual, most IM applications allow for the transmission of graphical symbols that convey emotions or emphasize the communication in some way, so-called "emoticons". Emoticons can for instance be a "Smiley" face, or a sad Smiley, or any graphical symbols that represents an emotion. ASCII text Smileys are well known from e-mail and text chat communication, and there are numerous variants that represent emotions. The emoticons in IM applications are graphical versions of those Smileys. Some IM applications provide means to generate various types of animation on the receiving end (e.g. to attract attention), and there are also examples of chat applications that have animated cartoon characters (avatars) where the Smileys (smiling face, sad face, etc.) are expressed on the character. All in all, emoticons, or emotional symbols, are used in many types of communication applications, and are also examples of a type of symbol that has multimodal implementations today (e.g. textual, graphical and animation, see example in figure 1). Such symbols could perhaps also benefit from being presented in other modalities (e.g. audio) in order to increase the communication experience for users.

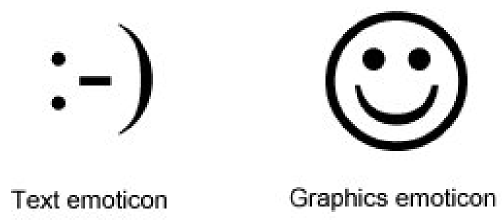

Figure C.2: Text and graphical version of smiling face emoticon

C.2.3 Real time communication services

Another area of communication interaction that is not represented in ISO/IEC 11581 is mobile communication and regular telephony. There are a number of symbols associated with telephony and text and voice messaging that should be considered. In traditional telephony, auditory information has always been extensively used for status indication (ringing, busy, disconnected, etc.). On display telephones (e.g. mobile telephones, ISDN telephones), these status indications have graphical, textual and sometimes vibration equivalents in addition (e.g. "incoming call alert"), although their implementations may vary between manufacturers. These symbols can be said to be current examples of multimodal symbols, in that they have textual, graphical, audible and vibration components.

With the increased penetration of handheld computers, the synchronization of data between a main computer and a handheld computer is an operation that is more and more common. Symbols representing objects or actions associated with such synchronization may be of importance here as well.

Annex D (informative): Examples of best practices for the application of multimodal symbols in selected ICT services and devices

Very often, the availability of excellent and relevant guidelines does not mean that designers use them. Actually, providing guidelines without appropriate tools for designers to apply them within their development paradigms is almost useless. Tools like checklist, automatic validators, training courses, etc. are always required to make the guidelines really usable.

Two of these are outlined in the following examples: the application of multimodal user interface components to a public vending machine, and the automatic validator of compliance of a web with WAI guidelines.

D.1 Ticket vending machine

Table D.1: Basic information about the MAE ticket-vending system

Title of the example:	MAE, a ticket-vending machine at the customer's service.
Provider	COMELTA (http://www.comelta.com/)
Location	Stations of Ferrocarrils de la Generalitat de Catalunya – Barcelona region – Spain Stations of Airport Railway – Stockholm - Sweden
Year of project completion	1997

The objective of the project was to develop a new ticket-vending machine to install in train and metro stations. The requirements of Design for All were applied to the design in order to include all the population as the potential users of the machine.

Based on a touch-screen system, MAE was designed to match to dimensional requirements for the whole population, including disabled people on wheel chair.

All the screens are voice-assisted, in four languages, to help users. It also contains a simplified menu based on marks around the screen-frame intended for the blind people (with Braille description). The navigation mode for blind people uses only four buttons (next, previous, accept and cancel), always located in the same position, that allows all the different operations of the machine.

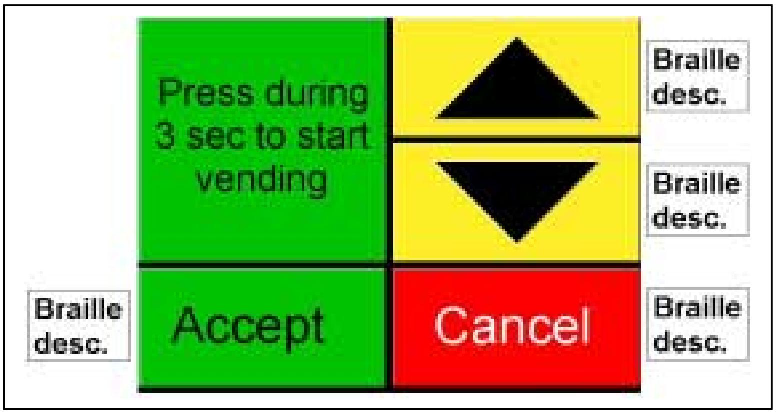

Figure D.1: Blind navigation screen

Figure D.2: First screen of the user interface of the application

Characteristics of the multimodal symbols in the project are:

- The first screen gives ten main options to the user: the nine tickets most often sold in every machine (according to each machine's memory) and an option for elderly people with direct and easy access to the range of low-prices fares intended to help them. By means of the first screen an 85 % of the ticket sale operations are made. It also presents an eleventh key which enables access to the rest of the transport titles and operations not included in the previous choices.
- Maximum visual contrast of colour of the elements concerning the environment and in all elements inside the display makes the visual detection easier.
- Buttons with 3D appearance and with acoustic feedback.
- The interface incorporates visual symbols in order to facilitate accessibility to all kind of people, especially for mentally disabled.
- All labels and text indications present appropriate size.
- Icons of "help" and "language change" are always available.
- User can choose the language of the different symbols on the screen.
- It has voice assisted screens both for the visually impaired and the foreign people.
- Tactile marks in the frame of the screen are the references needed to locate the options in order to activate the options of voice.
- All the operation buttons and slots have multimodal aspects to allow a simple interaction. Visual feedback (when an operation point is active and also images to recognize the functionality), acoustic feedback (voice help) and tactile feedback (each slot has a mark in Braille that explains its use).

The characteristics of the blind navigation mode are:

- When the system of voice is activated, the machine carries out a short process of training.
- The touch screen changes its appearance and it transforms in 4 buttons or sensitive areas.
- The activation of the buttons has two acoustic feedback modes: one at the moment of pressing and the other at the moment of releasing the button.
- All the information written in the screen is provided by spoken messages.

Access to navigation mode:

- To change from the standard interface to the navigation mode, a symbol at the bottom left of the screen must be clicked. As there were no standards on accessing to navigation mode in user interfaces (it was really a newly developed product) the solution in this case was designing this symbol. Through the Spanish organization for people with visual impairments, this information was added to the training of blind people in Spain using the underground.

- However, it is not an end solution. For example, foreign blind people probably will not be able to access to this navigation mode. For that reason, standardization in this issue is needed to adapt the interface to user profile (with an smart card, standard sequence in the screen, standard location for the navigation mode access button...).

D.2 Web accessibility validator

Table D.2: Basic information about the "Bobby" web site accessibility validator

Title of the example:	Bobby: a web-based tool that analyzes web pages for their accessibility to people.
Provider	CAST, Center for Applied Special Technology
Location	Internet: http://www.cast.org/bobby/
Year of project completion	1999-2000

Automatic validators

A validator can verify the syntax of your pages (e.g. HTML, CSS, XML). Correct syntax will help eliminate a number of accessibility problems since software can process well-formed documents more easily. Also, some validators can warn you of some accessibility problems based on syntax alone (e.g. a document is missing an attribute or property that is important to accessibility). Note, however, that correct syntax does not guarantee that a document will be accessible. For instance, you may provide a text equivalent for an image according to the language's specification, but the text may be inaccurate or insufficient. Some validators will therefore ask you questions and step you through more subjective parts of the analysis.

Some examples of automatic validators include:

- An automated accessibility validation tool (for example Bobby): it will help page authors identify necessary changes to their pages so users with disabilities can more easily use their Web pages. For example, a blind user will be aided by adding a sound track to a movie, and a hard-of-hearing user will be aided by a written transcript of a sound file on a Web page. Bobby will recommend that these be added if they do not already exist.

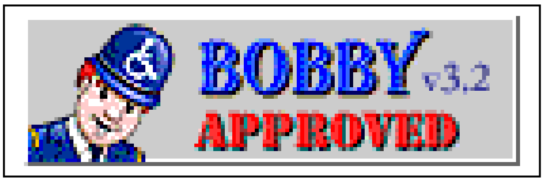

Figure D.3: Graphical image to indicate that the web site meets the web accessibility guidelines

Many people with disabilities will use special Web browsers, such as one that reads text out loud using a speech synthesizer for blind users. The suggestions made by Bobby will help authors to add information to a Web page that will help the special browsers work more effectively.

Other initiatives has arisen in different countries, as for example TAW (Test Accessibility of the Web) in Spain, with the same objective as Bobby.

- An HTML validation service (for example W3C HTML Validation Service).

- A style sheets validation service (for example W3C CSS Validation Service).

Human review

All these tests should highlight major access issues, and are valuable in reducing a number of accessibility barriers. Validators usually report what issues to solve and often give examples of how to solve them. They do not usually help an author walk through each problem and help the author modify the document interactively.

However, lots of accessibility aspects and requirements of the Web sites will not be discovered by the automatic tool, and always a human review is necessary. Some of the aspects related with symbols and icons that should be human-checked are:

- Identifying potentially sensitive cultural issues that might arise due to language or icon usage.
- Meaningfulness of link text and text equivalents.
- Check relation of text equivalent with auditory and visual content.
- Revision of colours used.
- Check that standardized icons and symbols are correctly used in the page.

Finally, we must have in mind that in real-life settings, your pages may be less usable than you expected (testing scenarios only replicate conditions caused by a disability; they do not simulate the full experience a user with a disability might have). Thus, one of the strategies recommends that content developers observe people with different disabilities as they attempt to use a page or site. Expert and novice users with disabilities will provide valuable feedback about accessibility or usability problems and their severity.

Annex E (informative): Useful information resources to design multimodal symbols

Abascal, J.G., Noirhomme-Fraiture, M., Petrie, H. and Van Nes, F. (1997). Guidelines for the design of HCI for people with disabilities. *Human-Computer Interaction: INTERACT '97.* (IFP TC13 International Conference on Human-Computer Interaction). London: Chapman and Hall. ISBN 0 412 80950 8.

"Bridging the gap?" COST 219bis 2001

CEN/CENELEC Guide 6: "Guidelines for standards developers to address the needs of older persons and persons with disability".

Colwell, C. and Petrie, H. (1999). A preliminary evaluation of the WAI guidelines for producing accessible web pages. In C. Bühler and H. Knops (Eds.), *Assistive technology on the threshold of the new millenium.* Amsterdam: IOS Press.

Davis, H.: "Abnormal hearing and deafness". In H.Davis and S.R.Siverman (eds). Holt, Rhinehart and Winston, New York

"Disability and social participation in Europe" Eurostat 2001

"Disabled persons statistical data" 2nd Ed. European Commission 1995

Edwards, A.D.N. and Holland, S. (1992), (eds): *Multimedia Interface Design in Education.* Berlin: Springer-Verlag.

ETR 334 (1996-12): "Human Factors (HF): "The implications of human ageing for the design of telephone terminals".

Fernstrom, M. and Bannon, L. (1997). Explorations in sonic browsing. In H. Thimbleby, B. O'Conaill and P. Thomas (eds) *People and Computers XII, Proceedings of HCI'97*. London: Springer

Fitch, W.T and Kramer, G. (1994). Sonifying the body electric: Superiority of an auditory over visual display in a complex, multivariate system. In G. Kramer (ed) *Auditory display: sonification, audification and auditory interfaces.* Reading, MA: Addison-Wesley.

Fredriksen, J. et al: "Use of telecommunications: the needs of people with disabilities" - Fundesco and Telefonica, Madrid, 1989.

Gaver, W.W. The Sonic Finder: (1989). An interface that uses auditory icons. *Human-Computer Interaction, 4,* pp 67 - 94.

Gaver, W.W. (1994). Using and creating auditory icons. In G. Kramer (ed*) Auditory display: sonification, audification and auditory interfaces*. Reading, MA: Addison-Wesley.

Gill, J.: "Which Button? Designing user interfaces for people with visual impairments" RNIB 2000

Gleiss, Norman: "The need of aids for hearing impaired telephone users" TELE No. 2, 1980

Glorig, Ward, Nixon: "Damage risk criteria and noise induced hearing loss". NPL conference on Control of Noise, 1961

Hannaford, B. & Venema, S. (1995): "Kinesthetic displays for remote and virtual environments", in W. Barfield and T. Furness (Eds.) *Virtual environments and advanced interface design*. New York: Oxford University Press.

Hardwick, A., Furner, S., Rush, J. (1998) Tactile display of virtual reality from the World Wide Web - a potential access method for blind people. *Displays Vol 18* Issue 3, pp 153-161 ISSN 01419382

Immersion TouchSense™ Fundamentals, Immersion Corporation 1998-2000, 801 Fox Lane, San Jose, CA 95131, Phone (408) 467-1900. Fax (408) 467-1901, Revision 1.2 help, October 25, 2000 (The latest version of the present document can be downloaded at: www.immersion.com)

Jameson, D.H. (1994): "Sonnet: audio-enhanced monitoring and debugging", in G. Kramer (ed) *Auditory display: sonification, audification and auditory interfaces*. Reading, MA: Addison-Wesley.

Johnson, V. and Petrie, H. (1999): "Requirements for the use of screenphones by older people". In C. Bühler and H. Knops (Eds.), *Assistive technology on the threshold of the new millenium.* Amsterdam: IOS Press.

Karshmer, A.I., Brawner, P. AND Reiswig, G. (1994): "An experimental sound-based hierarchical menu navigation system for visually handicapped use of graphical user interfaces". *In Proceedings of Assets '94: First Annual ACM Conference on Assistive Technologies*. New York: ACM Press.

Kember, P., Ainsworth, L. and Brightman, P.: "A hand anthropometric survey of British workers", Ergonomics Laboratory, Cranfield Institute of Technology, UK, 1981

Keyson D V (1996) Touch In User Interface Navigation, Thesis Technische Universiteit Eindhoven, ISBN 90-386-0288-X (available for download at: http://alexandria.tue.nl/extra3/proefschrift/PRF12B/9602316.pdf)

Loomis, J. M. and Lederman, S. J. (1986): "Tactual perception", In K. K. Boff, L. & Thomas, JP. (Eds.), *Handbook of perception and human performance*. New York: Wiley/Interscience.

Morley, S., Petrie, H., O'Neill, A.-M., and McNally, P. (1999): "Auditory navigation in hyperspace: design and evaluation of a non-visual hypermedia system for blind users", *Behaviour and Information Technology, 18*(1), pp 18 - 26.

Multimodal Interaction Styles, Bert Bongers, Berry Eggen, David Keyson and Steffen Pauws. Short paper, HCI'97 Conference Companion. *HCI'97 Conference on Human Computer Interaction*, Bristol, UK. August 1997.

Mynatt, E.D. (1997): "Transforming graphical interfaces Into auditory Interfaces for Blind Users", *Human-Computer Interaction, 12*, pp 7 - 45.

Mynatt, E.D. (1997): "Designing auditory cues for Mercator". In S. Howard, J. Hammond and G. Lingaard (eds*)* *Human-Computer Interaction: INTERACT '97*. (IFP TC13 International Conference on Human-Computer Interaction). London: Chapman and Hall. ISBN 0 412 80950 8.

Nielsen, J. (2000): "Is navigation useful?", available at http://www.useit.com/alertbox/20000109.html

Nielsen, J. (1995a): "Features for the next generation of web browsers". Available at: http://www.useit.com/alertbox/newfeatures.html

Nielsen, J. (1995b): "Navigating large information spaces". In J. Nielsen*: Multimedia and hypertext: the internet and beyond*. San Francisco: Morgan Kaufmann.

Nielsen, J. (1997): "The Tyranny of the Page: Continued Lack of Decent Navigation Support in Version 4 Browsers". Available at http://www.useit.com/alertbox/9711a.html

Nordby, K., Raanaas, R. K. and Magnussen, S. (2002): "The expanding telephone number. Part 1: Keying briefly presented multiple-digit numbers", *Behaviour & Information Technology* (in press) 12 pp.

O'Modhrain, S. & Brent, G. (1996): "Haptic user interfaces for the blind". Available at: http://ccrma-www.stanford.edu/CCRMA/Overview/node49.html

Petrie, H. (2000): "Acessibility and usability requirements for ICTs for disabled and elderly people: a functional classification approach". In J.G. Abascal and C. Nicolle (Eds.), *Inclusive guidelines for human computer interaction*. London: Taylor and Francis.

Petrie, H. and Carey, K. (2000): "Harnessing information technology: requirements for usability of information and communications technologies for people with disabilities". In *Disability Discrimination Act Access for All: a practical guide for professionals and business managers*. Swindon: British Computer Society

Petrie, H. and Morley, S. (1998): "The use of non-speech sounds in non-visual interfaces to the MS-Windows GUI for blind computer users". In A. Edwards and S. Brewster (Eds.), *Proceedings of the Fifth International Conference on Auditory Displays (ICAD '98).*

Petrie, H., Obee, J., Blenkhorn, P. and Evans, G. (1996): "User-based evaluation of a system for access to computer-aided software engineering (CASE) tools for blind software engineers. In J. Klaus, E. Auff, W. Kremser and W.L. Zagler (Eds.), *Interdisciplinary aspects on computers helping people with special needs*. Vienna: R. Oldenbourg. ISBN 3-7029-0413-1.

Pheasant, S.: "Bodyspace: Anthropometry, Ergonomics and design", Taylor and Francis, London, 1986.

Pitt, I.J. and Edwards, A.D.N. (1996): "Improving the usability of speech-based interfaces for blind user". *Proceedings of ASSETS '96: Second Annual ACM Conference on Assistive Technologies*. New York: ACM Press.

Raman, T.V. (1997*):* "Auditory user interfaces: toward the speaking computer", Boston: Kluwer Academic Publishers.

Raanaas, R. K., Nordby, K. and Magnussen, S. (2002). 'The expanding telephone number. Part 2: Age variations in immediate memory for multiple-digit numbers', *Behaviour & Information Technology* (in press) 7 pp.

Ramstein, C. (1996): "Combining haptic and braille technologies: design issues and pilot study", *Proceedings of ASSETS '96: Second Annual ACM Conference on Assistive Technologies*. New York: ACM Press.

Rigas, D.I. and Alty, J.L. (1997): "The use of music in a graphical interface for the visually impaired". In S. Howard, J. Hammond and G. Lingaard (eds) *Human-Computer Interaction: INTERACT '97*. (IFP TC13 International Conference on Human-Computer Interaction). London: Chapman and Hall. ISBN 0 412 80950 8.

Roe, Patrick R. W. (ed.) (2001):"Bridging the GAP? Access to telecommunications for all people". Published by the Commission of the European Communities/COST 219bis, (no ISBN given).

Rosenfeld, L. and Morville, P. (1998): "Information architecture on the World Wide Web: designing large scale web sites". O'Reilly.

Sikora, C.A. and Roberts, L.A. (1997): "Defining a family of feedback signals for multimedia communication devices". In S. Howard, J. Hammond and G.Lingaard (eds) *Human-Computer Interaction: INTERACT '97*. (IFP TC13 International Conference on Human-Computer Interaction). London: Chapman and Hall. ISBN 0 412 80950 8.

"Statistics in Focus: Theme 1 General statistics": – 4/99 Eurostat.

Stevens, R.D., Edwards, A.D.N. and Harling, P.A. (1997). Access to mathematics for visually disabled students through multimodal interaction. *Human-Computer Interaction, 12*, pp 47 - 92.

Timberlake, S. (2000): "The basics of navigation". Available at: http://www.efuse.com/Design/navigation.html

Welford, A.T., Norris, A.H. and Shock, N.W.: "Speed and accuracy of movement and their changes with age", *Acta Psychologica 30*, Attention and performance ii (pp3-15) 1969.

Wenzel, E.M. (1994): "Spatial sound and sonification". In G. Kramer (ed) *Auditory display: sonification, audification and auditory interfaces*. Reading, MA: Addison-Wesley.

Annex F (informative): Detailed Review on European and USA legislation about Design for All

The european approach

The Communication from the EC, entitled "Towards a Barrier Free Europe for People with Disabilities", of may 2000 presents an overview of the current european legislation on Design for All:

The inclusion of a general non-discrimination article covering inter alia disability in the Treaty of Amsterdam provides the basis for a crucial leap forward to promote equal rights for people with disabilities at EU level. Based on Article 13 of the EC Treaty, the European Commission has adopted on 26 November 1999 a comprehensive anti-discrimination package. From a disability perspective, the relevant part of this package consists of a proposal for a directive in the field of employment and occupation prohibiting discrimination on all grounds of discrimination listed in Article 13 and an action programme consisting of a wide array of complementary measures in this respect.

Other relevant initiatives have included:

- the Directive 98/10/EC on open network provision to voice telephony and on universal service for telecommunications in a competitive environment (98/10/EC OJ L 101, 01.04.1998) which requests the Member States, where appropriate, to take suitable measures in order to guarantee access to and affordability of all fixed public telephone services for disabled users and users with special social needs.

- More recently, on 9 March 1999 the Council and the European Parliament also adopted the Radio Equipment & Telecommunications Terminal Equipment (R&TTE) Directive (99/05/EC OJ L 091, 07.04.1999) which inter alia gives the Commission the powers to decide that apparatus within certain equipment classes or apparatus of particular types must be so constructed that it supports certain features in order to facilitate its use by users with a disability.

- In the frame of the European Information Society Standardisation policy, the Commission has also assigned a standardization mandate to the European standardization organizations, CEN, CENELEC and ETSI, inviting them to identify the specific needs for standardization in support of a better integration of elderly and disabled people in the Information Society (see http://www.cenorm.be/).

- Under the Fifth RTD Framework Programme (Decision No 182/1999/EC of the European Parliament and of the Council of 22 December 1998, OJ L 026 du 01.02.1999), research notably addresses person/system interfaces and adaptive and assistive systems to overcome problems caused by environmental barriers confronting those with physical and/or intellectual impairments, as well as intelligent systems and services to support independent living and participation in the information society.

- The European Commission's Joint Research Centre also actively participates in research related to a number of emerging technologies; such as the implementation of voice-to- text systems for the deaf, man-machine interfaces at the workplace, and internet access for people with disabilities.

This positive trend is also been reflected in the development of the European consumer policy which aims to cover all kinds of consumers including consumers with disabilities. Examples of this type of action include the development of criteria on Special User Needs that could be included in the protocols used for comparative testing of consumer products and the application of the principle of "Design for all"; although the latter has to date concentrated primarily on preparation for the introduction of the EURO. This work has been carried out through the integration of representatives of people with particular impairments at all stages of the decision making process as consumer representatives.

In the context of consumer protection, mention should also be made of the Directive 95/46/EC 21 which ensures inter alia a qualified protection of personal data revealing disabilities. (21 OJ L 281, 23.11.1995.)

In general, the european approach to legislate on Design for All matters has been based on promoting consensus between countries, on funding research initiatives and issuing mandates to relevant technical and organizational bodies.

The USA approach

Over the past 25 years the USA has produced an array of laws and regulations relating to IT and disability:

The Americans with Disabilities Act (ADA)

The Americans with Disabilities Act (ADA) of 1990 mandated that public services, places of public accommodation, and telecommunications services be accessible to citizens with disabilities. The ADA prohibits discrimination against people with disabilities in "the full and equal enjoyment of goods, services, facilities, privileges, advantages, or accommodations of any place of public accommodation". Public places such as hotels, restaurants, and stores are required to provide auxiliary aids and services and to remove architectural barriers. The requirements of the ADA include requirements that information provided to the general public (for example via electronic kiosks) also be accessible to people with disabilities. This requirement is not restricted to removing architectural barriers for the mobility impaired, rather, it includes a the accessibility of information to individuals with sensory and cognitive impairments.

The ADA requires that employers provide "reasonable accommodation" to employees with disabilities. Employers must make modifications or adjustments to the job application process, modifications or adjustments to the work environment, and modifications or adjustments that enable an employee with a disability to enjoy equal benefits and privileges of employment.

The ADA also requires that state and local governments take steps to ensure communications with applicants, participants, and members of the public with disabilities are as effective as communications with others

The Telecommunications Act of 1996

This law contains the guidelines that are required to principally address the access needs of individuals with disabilities affecting hearing, vision, movement, manipulation, speech, and interpretation of information.

Section 255 provides that a manufacturer of telecommunications equipment or customer premises equipment shall ensure that the equipment is designed, developed, and fabricated to be accessible to and usable by individuals with disabilities**, if readily achievable**. A provider of telecommunications services shall ensure that the service is accessible to and usable by individuals with disabilities, if readily achievable. Whenever either of these is not readily achievable, a manufacturer or provider shall ensure that the equipment or service is compatible with existing peripheral devices or specialized customer premises equipment commonly used by individuals with disabilities to achieve access, if readily achievable.

Especially interesting sections in this law are:

Section 1193.23 Product design, development and evaluation: this section requires manufacturers to evaluate the accessibility, usability, and compatibility of telecommunications equipment and customer premises equipment and incorporate such evaluation throughout product design, development, and fabrication, as early and consistently as possible. Manufacturers must develop a process to ensure that products are designed, developed and fabricated to be accessible whenever it is readily achievable. Since what is readily achievable will vary according to the stage of development (i.e. some things will be readily achievable in the design phase which are not in later phases), barriers to accessibility, usability, and compatibility must be identified throughout product design and development, from conceptualization to production. Moreover, usability can be seriously affected even after production, if information is not provided in an effective manner.

The details of such a process will vary from one company to the next, so this section does not specify the structure or specific content of a process. Instead, this section sets forth a series of factors that a manufacturer must consider in developing such a process. How, and to what extent, each of the factors is incorporated in a specific process is up to the manufacturer.

Section 1193.31 Accessibility and usability: this section provides that, subject to section 1193.21, manufacturers must design, develop and fabricate their products to meet the specific requirements of sections 1193.33 through 1193.43. As discussed under section 1193.21, some sections related to usability have been moved to this subpart to reflect that they are subject to the readily achievable limitation. The title has been changed and the sections renumbered accordingly.

And there are also interesting "Frequently Asked Questions" (FAQ) about this Act:

Q: What are manufacturers required to do?

A: Manufacturers must ensure that such equipment is accessible and usable, if it is readily achievable. If accessibility is not readily achievable, the manufacturer must make the equipment compatible with peripheral devices used by people with disabilities, if that is readily achievable.

Q: How will the requirements be enforced?

A: Complaints can be filed with the Federal Communications Commission (FCC), which has sole jurisdiction over enforcement; there is no "private right of action."

Q: What does "readily achievable" mean?

A: Readily achievable has the same meaning as in the ADA: easily accomplishable without much difficulty or expense.

Q: Who decides whether something is readily achievable?

A: The manufacturer, based on the cost and its resources.

A number of other federal regulations require consumer electronics be designed to include the disabled:

- The Telecommunications for the Disabled Act of 1982 requires that all "essential telephones" be hearing aid compatible.
- The Hearing Aid Compatibility Act of 1988 requires that all wireline telephones manufactured in the US or imported for use in the US after August 16, 1989 be hearing aid compatible.
- Cordless telephones also were required to be hearing aid compatible by August 16, 1991.
- The Telecommunications Accessibility Enhancement Act of 1988 requires the Federal telecommunications system to be fully accessible to individuals with hearing and speech disabilities.
- Section 508 of the Rehabilitation Act, as amended in 1991 and 1992, requires that the General Services Administration develop and adopt guidelines to ensure federal employees with disabilities can use electronic office equipment and information technologies (leased or purchased) with or without special peripherals.

The section 508 of the Telecommunications Act and the Information Society: there is a similar trend for accessibility of information including information infrastructures such as the Internet. Many requirements (such as video captioning) are expected to be required of information available to the public through the Internet. Mass market consumer products that meet the needs of both the average individual and the needs of the disabled will be best positioned to serve these markets.

Under standards published by the Board on December 21, 2000, the Federal government will be in the forefront in ensuring access to electronic and information technology. These standards cover various means of disseminating information, including computers, software, and electronic office equipment. They provide criteria that spell out what makes these products accessible to people with disabilities, including those with vision, hearing, and mobility impairments. The Board developed these standards under section 508 of the Rehabilitation Act as amended by Congress in 1998. The law applies to all Federal agencies when they develop, procure, maintain, or use such technology. Federal agencies must ensure that this technology is accessible to employees and members of the public with disabilities to the extent it does not pose an "undue burden." The law directed the Board to develop access standards that are to become part of the Federal government's procurement regulations. The scope of section 508 and the Board's standards are limited to the Federal sector.

The new standards provide technical criteria specific to various types of technologies and performance-based requirements, which focus on the functional capabilities of covered technologies. Specific criteria cover software applications and operating systems; web-based information or applications; telecommunications functions; video or multi-media products; self contained, closed products such as information kiosks and transaction machines, and computers. Also covered is compatibility with adaptive equipment people with disabilities commonly use for information and communication access.

The standards are based on recommendations from an advisory committee the Board established for this purpose. The Electronic and Information Technology Access Advisory Committee was composed of 27 members representing industry, various disability organizations, and other groups with an interest in the issues to be addressed. The Board published the standards in proposed form on March 31, 2000 and made them available for public comment for 60 days. Over 100 individuals and organizations submitted comments on the standards. Comments were submitted by Federal agencies, representatives of the information technology industry, disability groups, and persons with disabilities. The Board finalized the standards according to its review and analysis of these comments.

USA Federal government has also requested private industry and research entities to provide significant initiative in improving access for people with disabilities in the Digital Age. These initiatives will include: ensuring that computer scientists and engineers receive training on accessibility; expanding the number of faculty who conduct research on accessibility; and ensuring that university online resources are accessible to people with disabilities.

USA Administration Agencies in charge of accessibility issues

Furthermore, the United States has created specific organizations in its administration to deal with accessibility matters, the Access Board. This federal agency has the following committments:

- develops minimum guidelines and requirements for standards issued under the Americans with Disabilities Act (ADA) and the Architectural Barriers Act (ABA);
- develops accessibility guidelines for telecommunications equipment and customer premises equipment under the Telecommunications Act;
- develops accessibility standards for electronic and information technology under section 508 of the Rehabilitation Act;
- provides technical assistance on those guidelines and standards;
- and enforces the Architectural Barriers Act.

A Committee was convened by the Access Board in June 1996 to assist the Board in fulfilling its mandate to issue guidelines under the Telecommunications Act. The Committee was composed of representatives of manufacturers of telecommunications equipment and customer premises equipment; manufacturers of specialized customer premises equipment and peripheral devices; manufacturers of software; organizations representing the access needs of individuals with disabilities; telecommunications providers and carriers; and other persons affected by the guidelines. Also a Telecommunications Access Advisory Committee (TAAC) has been formed to address these issues. The EIA, CEMA, and the Telecommunications Industries Association (TIA) are represented on this commitee.

The Access Board and the General Services Administration (GSA) are directed to provide technical assistance to individuals and Federal agencies concerning the requirements of section 508. The Federal Information Technology Accessibility Initiative (FITAI) is an interagency effort, coordinated by GSA, to offer technical assistance and to provide an informal means of cooperation and sharing of information on implementation of section 508. Under a contract awarded to a private firm in September, the Board is developing training modules and technical assistance materials on the new standards and section 508. Materials to be developed include fact sheets, brochures, answers to frequently asked questions, multimedia presentations, narrated slide shows, practical "how-to" tips on making web sites accessible to people with disabilities, and comprehensive annotated lists of reference materials. Training modules will be available for use by speakers at workshops and conferences. The modules will be designed to meet the needs of various audiences, such as Federal managers, end-users with disabilities, the procurement community, and producers of technology.

Questions about the 508 standards should be sent to 508@access-board.gov.

There are several important differences:

- The EC has promoted consensus and issued general directives, while the US have made concrete laws addressing the problem. This effectively has created a broad mandate for accessibility in the US, while in Europe there is still the need to harmonize legislation and standards.
- Rather that addressing inaccessible technology already in place, US law addresses technologies while they are still under development. For example, the Telecommunications Reform Act (1996) requires all telecommunication manufacturers and service providers to ensure that new products are accessible or compatible with assistive technology where readily achievable.

The EC recognizes the effectiveness of the US approach in the "Mandate to the European Standards Bodies for Standardization in the field of information and communications technologies (ICT) for disabled and elderly people" Brussels, 6 May 1998 DG XIII-C5/DR/D(97)" with the following words:

- "A likely effect of this legislation will be that US industry will be better placed to respond to the demographic changes in society and corresponding changes in the market."

Annex G (informative): Bibliography

ISO/IEC 11581-4 (2000): "Information technology - User system interfaces and symbols - Icon symbols and functions - Part 4: Control icons".

ISO/IEC 11581-5 (2000): "Information technology - User system interfaces and symbols - Icon symbols and functions - Part 5: Tool icons".

ISO 14915 (Draft): "Software ergonomics for multimedia user interfaces".

DigiScents (2001): "DigiScents - A Revolution of the Senses", http://www.digiscents.com/ [25 September 2001].

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

History

Document history		
V1.1.1	June 2002	Membership Approval Procedure MV 20020823: 2002-06-25 to 2002-08-23