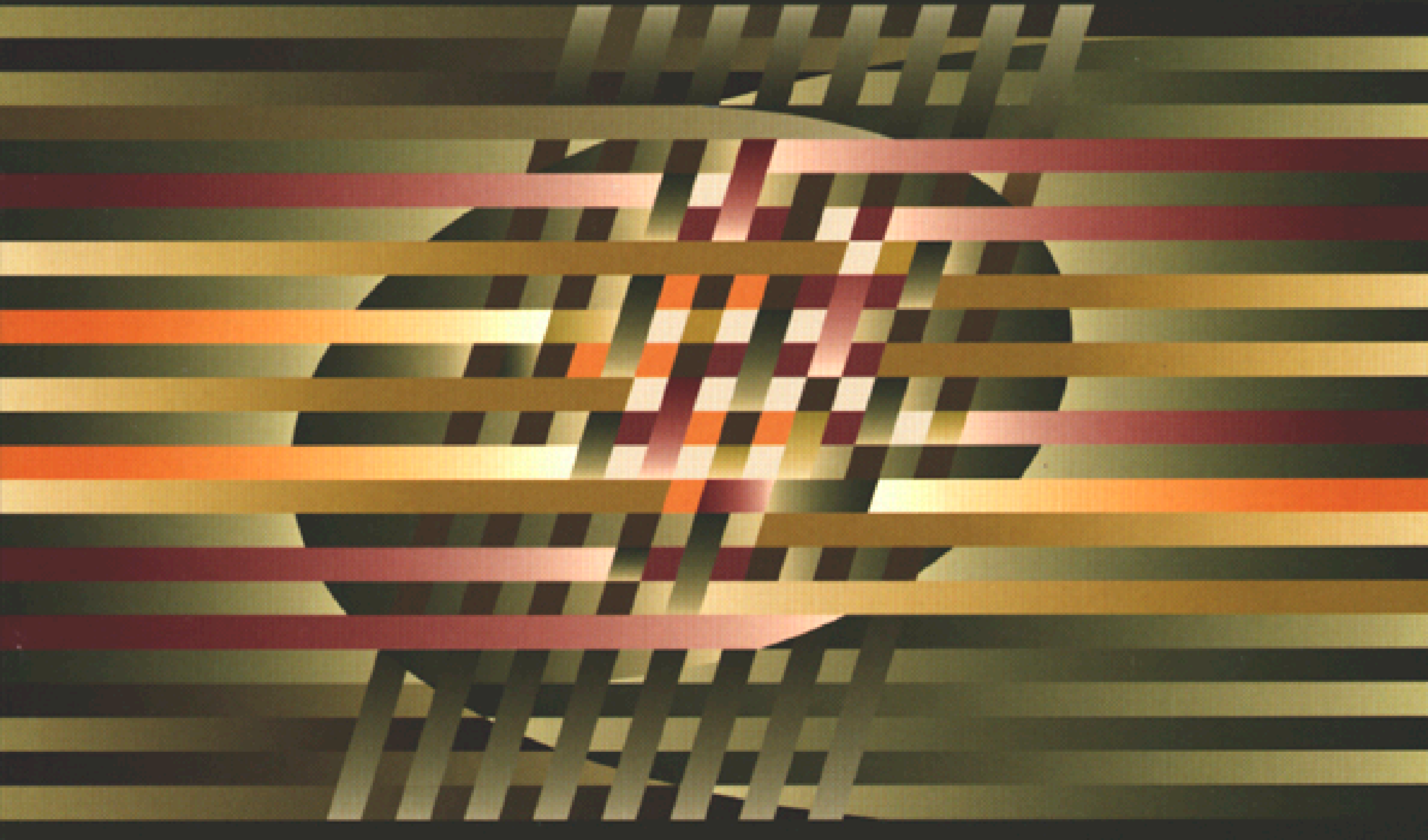
Telektronikk 3/4.97

Human Factors in telecommunication



Contents

Feature

Special

Status

Telektronikk

Volume 93 No. 3/4 - 1997
ISSN 0085-7130

Editor:
Ola Espvik
Tel. + 47 63 84 88 83

Status section editor:
Per Hjalmar Lehne
Tel. + 47 63 84 88 26

Editorial assistant:
Gunhild Luke
Tel. + 47 63 84 86 52

Telenor

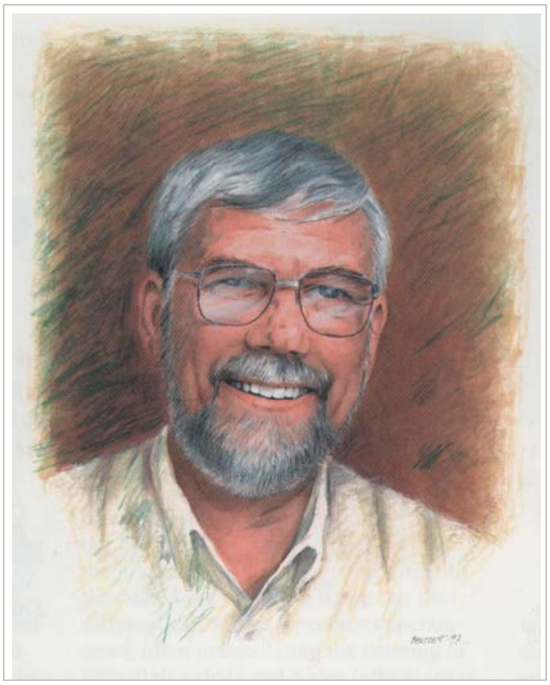
Editorial office:
Telektronikk
Telenor AS, Telenor Research & Development
P.O. Box 83
N-2007 Kjeller, Norway
email: telektronikk@fou.telenor.no

Graphic design:
Design Consult AS

Layout and illustrations:
Gunhild Luke, Britt Kjus, Åse Aardal
Telenor Research & Development

Guest editorial

KNUT NORDBY



The possible services that can be offered by today's telecommunication systems are staggering, and new services are continually being added. However, developers and operators often seem to forget that the end users of these sophisticated services belong to the species *Homo sapiens*, which has not changed significantly during the last 50 000 years. Although today's computer-controlled telecommunication networks should be able to offer services fully adapted to the capabilities of human end users better than any previous technology, the opposite is usually the case and new services that are not specifically adapted to human capabilities are regularly introduced. Thus, the user interface of the plain old telephone (POT) has actually deteriorated from the time of Alexander Graham Bell to some of today's interactive voice services. In those early days users interacted with other people and any misunderstandings could be set straight right away using the most basic of all human capabilities – the spoken language.

If people cannot easily use technology we will increasingly see that the people in the Information Society will be divided into the 'information rich' and the 'information poor'. The former group will successively improve its position in society, while the latter may be excluded from important societal services and become increasingly dependent on others. More than 25 % of all European citizens are either aged, disabled or both, and these do not in any way make up any marginal groups.

During the last ten years a wave of liberalisation, de-regulation and privatisation of the state-owned telecommunication monopolies has swept across Europe and other parts of the world. These liberalised, privatised telecommunication markets need new regulation, and a wave of re-regulation and legislation is now under way: the most liberalised market of them all, the US market, is also the most regulated market to secure basic services, just trading and fair competition.

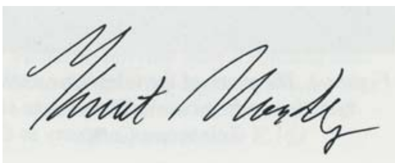
It is to these re-regulated markets that user-groups, special interest groups and consumers' organisations are moving in, demanding proper services and equipment. The trade journals now print comparative tests of similar products, and these tests review usability; i.e. how easy is it to use, how long does it take to learn to use it, is it easy to make a fatal mistake. And purchasing decisions are increasingly made on the basis of such reviews. Prospective customers do not only look at prices – ease of use has become an important decision factor – a product with a bad usability mark is not going to make it in the marketplace. Thus, manufacturers and service providers are going to need all the human factors expertise they can get to keep a competitive advantage.

On the other side we have the legislators. Society cannot sit still and watch large groups of citizens being left outside the information society. Again, the US is leading. In 1990 Congress passed the *Americans with Disabilities Act* (ADA), and a new *Telecommunications Act* is now going through Congress. This legislation provides for extensive rights for the users and is being closely studied by other countries. In Europe, there is an increasing demand for Pan-European legislation and a mounting pressure on the Commission of the European Communities (CEC) to issue Directives to achieve the same level of consumer protection that US citizens enjoy.

In the days of the state-controlled monopolies, the consumers had no way to influence the quality and usability of services and equipment. Today, consumers' and end-users' organisations participate in standardisation, thus filling a void left by telecom operators' who are reluctant to partake in standardisation – and these consumers' organisations may soon be the ones who are calling the shots. Actually, the service providers and the consumers have the same objectives – good usability – but they may have different solutions as how to achieve them.

For various reasons industry, and the telecom industry is no exception, has come to regard good usability or the misused term 'user friendliness' as costly add-ons that offer little advantage in the long run. Today, this view is being turned on its head. Simple cost/benefit analyses will show even the most stubborn executives the great economical and commercial advantage of good usability: When NYNEX, the great US East-coast based telephone company, managed to shave just *three seconds* average off each inquiry transaction, they saved themselves hiring more than 500 operators.

The bottom line of all this is *Human Factors* or *Ergonomics*, the science of designing good user interfaces. The following nine articles in this *Feature Section* of *Telektronikk* are devoted to various perspectives of Human Factors in Telecommunications. These range from simple tactile identifiers on telephone cards to empowering users of new technology; from sound quality in the telephone to the special requirements of tele-working; from the basic tones (line signals) we hear in the handset to the intricacies of using video-telephones, and from the historic origins of human factors in telephony (did Alexander Graham Bell think about it?) to a review of contemporary issues. A brief overview like this can only scratch the surface of a large field, but hopefully this may inspire some readers to dip deeper into this fascinating topic.



Do we need human factors in telecommunications?

KNUT NORDBY

The changing user interface

When the telephone was in its infancy, the user interface was the simplest of all times: You just lifted the handset off-hook and gave a small crank a turn or two to alert the other party, and when the call was answered you could start the conversation. This was in the days when only two and two telephones were connected together (e.g. between the home and the shop, or between the mansion house and the porter's lodge). However, the idea of connecting several telephones together via a *switchboard* or *exchange* was soon realised, making it possible to reach many people, thus making the telephone a much more important communication instrument. You still only needed to lift the handset off hook and give the little crank a turn or two to alert the *operator* (in some systems you had to press the hook a couple of times), and when the exchange answered you simply asked the operator to connect you to the party you wished to speak to. When the call was finished, you merely replaced the handset on-hook (in some systems you had to give the crank another turn to alert the operator to disconnect you) – in other words, a simple, intuitive, speech based user interface which could be used by anyone who could speak the language (see Figures 1 and 2).

NORWEGIAN TELECOMMUNICATIONS MUSEUM

Figure 1 Norwegian telephone set from 1893 with inductor crank. This telephone set worked with the simple user interface from the early days of telephony. The earpiece on the handset can be moved to adapt to 'small and large heads', a simple ergonomic consession to user variability

NORWEGIAN TELECOMMUNICATIONS MUSEUM

Figure 3 Norwegian telephone set from 1931 with rotating number dial. This was the first 'modern' telephone with integrated hook switch. It was designed by Johan Christian Bjerknes of Elektrisk Bureau, Norway (with some finishing touches by the artist Jean Heiberg), but it was to be produced in Sweden by L.M. Ericsson, who had acquired a majority of Elektrisk Bureau shares

As the telephone networks grew, operators could no longer keep track of the names and numbers of all subscribers, and the telephone companies started issuing printed directories listing their subscribers' names and numbers. Before making a call you now had to look up the number of the party you wished to speak to in the directory and communicate it to the operator – the *telephone number* and the *telephone book* had arrived. In the beginning, this was not a serious problem for most users, since the subscriber lists were relatively short, phone numbers had only three or four digits and the print was large and legible. As the number of subscribers increased, phone books grew bigger and heavier, while the print was made smaller on low quality paper to limit the size of the telephone book to keep costs down. This soon created problems, especially for elderly people with failing eye-sight and for visually impaired people – a problem that persists to this day [1,2].

NORWEGIAN TELECOMMUNICATIONS MUSEUM

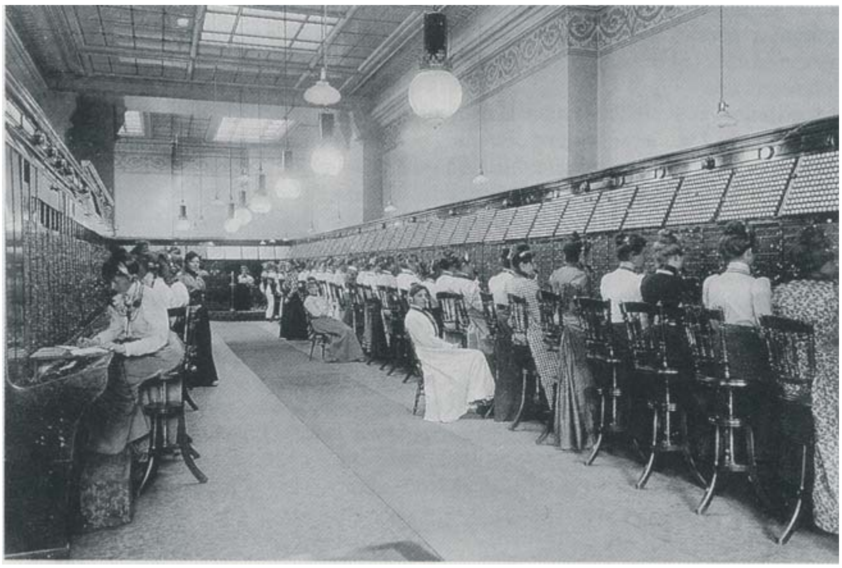


Figure 2 The work of the telephone exchange operators (nearly always women) was a very demanding chore. The picture shows the main exchange of the Christiania Telephone Company in Christiania (now Oslo) in 1896

This was only the first of many turns for the worse in the telephone user interface. It was, of course, still possible to ask the operator to look up an unknown number, but this practice had to be dropped as traffic grew and the work load on operators increased. Instead directory services were introduced: Special numbers were established that you could call to access directory information. This information was, and still is, given by operators, with whom you can talk and solve misunderstandings. Initially, this service was free, but, eventually, it was turned into a revenue maker. For people with failing eyesight or who are blind and who cannot look up numbers in a directory this has now become an economic problem since the prices for directory services have gone up, and alternatives (directories on CD-ROMs or diskettes) are very expensive or not available. In some countries (e.g. Norway) blind subscribers are given a limited number of directory enquiries free, but these have to be made from the home (or work) telephone. However, it is usually when you are on the move when you have no directory available that you need the service most.

The introduction of the automatic telephone exchange and telephones with turning number dials was a momentous setback for the simple user interface. In addition to having to look up the number in the directory, you now also had to be able to listen for, discriminate between and understand several line signals (the various tones you hear in the handset, such as *dialling tone, ringing tone, engaged tone*). But first and foremost, you had to 'send' the telephone number to the exchange yourself. This was achieved by inserting the index finger in the hole at the desired digit in the telephone dial and turn the dial clockwise to the stop, withdraw the finger to let the dial run back by itself, and then repeat this procedure until all the digits in the telephone number had been dialled. The movements required to operate a turning telephone dial are hard enough for young healthy people, but for elderly people with rheumatism or arthritis, or for people with upper limb motor and dexterity impairments, operating the turning dial can be a serious or even insurmountable obstacle (see Figure 3).

For operators at telephone exchanges, the physical strain of dialling telephone numbers all day was so great that various devices had to be introduced to lessen their nasty chore. One such device (see Figure 4) was a push-key number sender, a device with ten typewriter-like push-keys numbered 1 through 9 + 0, to ease the heavy load imposed on index finger, hand, wrist and arm; anticipating the push-button telephone by several decades. However, push-key diallers were normally not available to the ordinary subscriber, who would still have to struggle with the turning dial for many years. This should, by the way, remind us that good user interfaces and improved usability is not only an issue for the subscriber or end user, but is just as important for all people working to provide the service. Substantial savings have been made by telephone companies by improving system interfaces and the operators' work environment.

The cognitive problems of setting up calls, with no feedback other than the line signal tones to tell you where in the process you are, can be a daunting task even for resourceful people, but may be impossible to perform by people with learning disabilities or cognitive impairment. Thus, many elderly or mentally retarded people may have to receive help from another person to make a telephone call [1,2].

At first, direct dialling was only available for local calls, while long distance and international calls still had to be handled by operators. But by leaps and bounds direct dialling 'subscriber trunk dialling' has been extended, first for long distance domestic calls, and eventually also for international calls. It is possible today to call virtually any telephone in the world directly from your own telephone without the intervention of an operator, but the price for this freedom (in addition to the cost of the call) is to figure out the different call set-up procedures encountered, often necessitating the entering of long digit strings and codes (which today may run into 11–12 digits) and the need to be able to discriminate between and understand many strange and unknown line signals or tones – at the last count Europe alone had 228 different line signals! [3,4,5,6].

The introduction of tone signalling (DTMF = Dual Tone Modulation Frequencies), touch-tone to some, and the replacement of the rotating dial with the now ubiquitous twelve-key numeric push-button key-pad has been a significant step forward for all users: To press keys marked with digits are both physically easier and cognitively simpler than turning a dial (see Figure 5).

The advent of interactive stored voice supplementary services, however, involving the use of the new ***** and **#** keys (referred to as *star* or *asterisk,* and *square, hash* [UK] or *pound* [US], respectively), have left many users dumbfounded. There is an urgent need for standardisation and guidelines on the use of these two symbols as there are no standardised rules or recommendations for their use today, and neither ITU-T (International Telecommunication Union, founded more than 125 years ago) nor ETSI (European Telecommunication Standards Institute, founded ten years ago) have made any serious attempts to rectify this situation by issuing harmonised recommendations or standards [7,8,9].

Even when standards and recommendations are available, however, these are

KNUT NORDBY

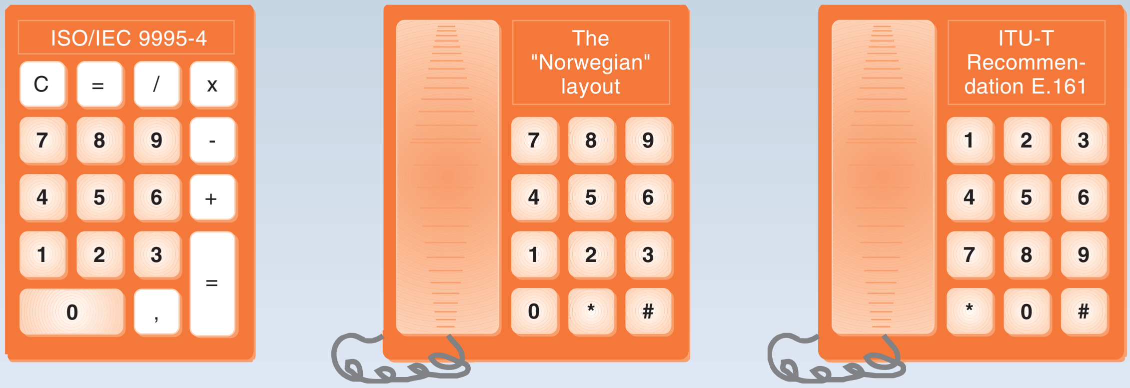
Figure 4 British (General Post Office) push key number dialler. This device was used in exchanges to alleviate the physical strain on operators who had to dial telephone numbers all day on a rotating dial

not always followed, often with serious consequences for the users – as demonstrated in the case of layout of digits and symbols on telephone key-pads in Norway. According to ITU-T Recommendation E.161 [10] the layout of digits and symbols on telephone dialling key-pads shall be (as shown in separate box, right panel) with **1 2 3** on the upper row and *** 0 #** at the bottom row. Norway, however, when tone signalling and push-button telephones were introduced in 1981, adopted a layout based on the International Organization for Standardisation, ISO, 9995-4 [11] for the layout of digits on calculator key-pads, with **7 8 9** at the top and **0 * #** at the bottom (as shown in separate box, middle and right panels). This was a well intended, but misguided,

NORWEGIAN TELECOMMUNICATIONS MUSEUM

Figure 5 'Tastafon', the first Norwegian 'touch-tone' push-button (DTMF) telephone from 1981. Push buttons are much easier to operate than a rotating dial. Note the inverted layout of the number keys which does not follow ITU-T Recommendation E.161

Three different digit layouts of push button keypads used in Norway: Left, the ISO/IEC 9995-4 standard layout used on calculators and PC keyboards [11]. Centre, the 'Norwegian' layout, based on the calculator layout. Right, the international ITU-T, Recommendation E.161 standard layout of telephone dialling keys [10]. When the push button 'touch-tone' telephone was being developed in Norway, early prototypes carried the international CCITT (now ITU-T) standard digit layout (right). However, when the push button 'Tastafon' telephones were introduced in 1981, this had been changed to the 'Norwegian' type layout (centre) based on the calculator type layout (left) "to aid the users by only having to interact with one common digit layout on telephones and calculators". Today, both the 'Norwegian' and the ITU-T layouts are used concurrently on telephones in Norway, making it nearly impossible for blind and visually impaired people to dial numbers correctly when encountering a strange telephone, and creating a general nuisance to all users who have to switch between the two layouts [10,11].



attempt to aid users by making telephone and calculator keypads similar, but which has turned out to be an encumbrance to nearly all users now that telephones with ITU type key-pad layouts can be sold alongside telephones with calculator type key-pad layouts. Both visitors to Norway and Norwegians travelling abroad encounter problems when switching between layouts, but those who are most affected are blind and visually impaired people. In Norway they have no way of knowing which way the digits are arranged when encountering strange telephones they have to use [12].

The concurrent use of different digit layouts on telephones was not new to Norwegians, who for many years had to cope with two different telephone dial layouts in Norway and a third layout in neighbouring Sweden (se separate box).

Putting a *tactile marker* (e.g. a raised dot or dimple, which can be felt by the sense of touch) on the **5** key does not help since the **4 5 6** keys are at the same positions (i.e. the second row) in both key layouts (also the **#** key has the same position in both layouts). Initially, Denmark also used the calculator type layout on telephones, but soon realised the mistake and changed to the ITU type layout after only a few years. People in both countries, however, will be hampered by the two concurrent layouts for many years to come since telephones seem to be rather long-lived instruments [10,11,12].

Telephone numbers are increasingly growing longer – many countries have already adopted nine-digit internal telephone numbers (e.g. France and United Kingdom) and there are plans for a Pan-European numbering system running into 13 digits, including the new four-digit operator codes (e.g. 1531 for Telenor). Many of the new interactive stored voice services (e.g. tele-banking, tele-shopping), and UPT (Universal Personal Telephone, called Alpha number in Norway) require entering long digit strings, including service codes, account numbers and PIN-codes (Personal Identification Number) – a development which Donald Norman has referred to as "The conspiracy against the human memory" [13,14].

System feedback to users of Audiotex services is usually limited to a mindless 'beep' or at best to a courteous, but unessential, voice message such as the obligatory "Thank you for your order" to indicate that a *valid* action has been made, but feedback is seldom given to confirm if the *desired* action has been correctly made [7,8,9].

The users must therefore devise their own procedures and strategies to be able to remember and dial the long digit strings required, such as writing down numbers, splitting long numbers into groups (chunks), or using pre-programmed automatic dialling devices. To avoid the many mistakes users invariably will make as numbers grow longer, *en-bloc dialling*, where the number is first composed on a display in the telephone and then sent as a whole block to the exchange by activating a "Send" button, (as with GSM and ISDN phones today), will have to replace traditional *overlap dialling* (where each digit is sent on to the exchange as it is dialled).

Human Factors in telecommunications

It was inevitable that the need arose for specialists to participate in the design of improved user interfaces. In 1947 Bell Labs hired psychologist John Karlin to do exactly this, and the discipline *human factors in telecommunications* was born. In the 50 years which have passed, the number of human factors specialists, most of them psychologists, have risen to over 200 at AT&T alone, and most leading telecommunication research laboratories around the world have seen similar development. Human factors as such did not arise in 1947, but some years earlier. Actually, its roots go back a good number of years to the emergence of industrial hygiene and industrial safety, and to ergonomics based on antropometric measurements (the measurement of man and human abilities).

The concept of *human factors* (mainly US) or *ergonomics* (mainly UK) as a distinct discipline did not arise within telecommunications, but with the military during World War II when it became evident that the physical, cognitive and mental abilities and the reactions of human operators in stressed and unexpected situations had to be taken into account when designing equipment and systems. It got an impetus after a number of freak accidents as pilots, having flown dangerous missions and returned unharmed to their bases, unexplainably crashed their planes after making some incomprehensible action, such as opening the bomb bay instead of extending the landing gear or similar slips. From its military uses, human factors thinking was taken up by various civilian fields such as civil aviation, industrial safety and eventually, telecommunications. At first, the interests in telecommunications were mostly concerned with speech quality and traditional telephony, but has now moved into virtually every field of telecommunications and information technology (IT), from numbering plans to supplementary services, from Internet telephony to tele-commuting.

The third root of human factors is more recent, *viz.* the advent of the electronic digital computer. At first, the big mainframes were operated by specially trained staff, but with the advent of the microcomputer (the PC, and not least the Mac) computers got into the hands of all kinds of people. In the beginning, cryptic commands and inscrutable operating systems created great problems for ordinary people using this new technology. This gave rise to HCI (Human Computer Interaction) and the HCI research led to the graphical user interfaces (GUI) with windows, menus, buttons, icons, pictograms, and object manipulation interaction. This started in the late 70s, but gained momentum during the 80s and is now a well established discipline. As telecommunications and IT are merging, so are also the fields of human factors in telecommunications and HCI until they now have become inseparable.

Why human factors?

Human factors is based on thorough in-depth knowledge about the physical and mental (i.e. cognitive) abilities of human users to address various areas in telecommunication. This may range from specification of new services to improving existing services, from user testing to writing style guides, from synthetic speech and speech recognition to design of web-sites, from the psychology of learning to organisational processes and so on, requiring cross-sector qualifications and excellence far beyond basic psychology [13,14].

MIKAEL JONSSON/MIRA/SAMFOTO

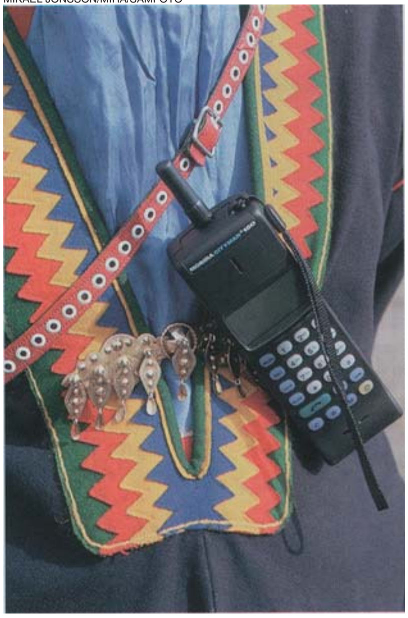


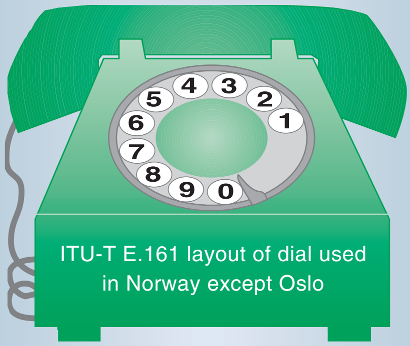
Figure 6 The success of the mobile telephone has been substantial and has spread to all sectors of society (here we see a Lapp user) despite the rather complicated user interface. If a product is felt to be very desirable or very practical, users are wont to put up with rather mediocre to bad user interfaces, but when the product has received a wide acceptance, demands for better usability soon arises

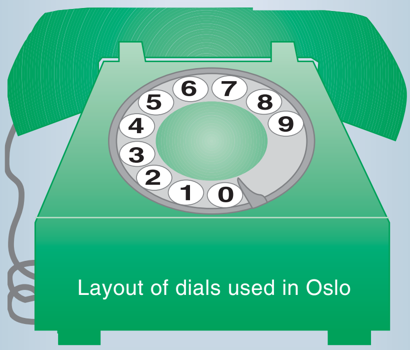
As deregulation, liberalisation and privatisation of the telecommunication sector has swept across Europe, the new-found competition has led to the prices, which were previously determined by political rather than market factors, now coming down. As prices are further pressed down by the increasing competition, operating companies have to look to other areas to gain a competitive edge. And the most current area here is usability, i.e. how user-friendly a service or product is.

We are long past the days when young computer buffs found a challenge in every new service or way of getting around in what we now call the Internet, trying out modems and new ways of getting around obstacles in the net. Users today demand fully 'transparent' services, i.e. they do not want (and should not need) to concern themselves with how a connection is established, whether terrestrial or by satellite, whether by analogue or digital transmission, whether line switched or packet switched,

Three telephone dial layouts

As a result of ordering equipment from producers who followed different technical standards, the digits on telephone dials in Oslo and its environs (centre) were laid out in the opposite order than on telephone dials in the rest of Norway (top), which comply to ITU Recommendation E.161 [10]. This created a great problem for blind people and others when travelling between Oslo and the rest of the country. The close neighbourhood with Sweden created an additional problem since Sweden used yet another non-standard layout (bottom). However, blind users were not completely at a loss since they always could rely on geography to guide them as to which dial layout to expect. With the two numeric key-pad layouts now in concurrent use this is not possible as there is no relation between layout and geography; it is quite possible to encounter telephones with both layouts even on the same desk.





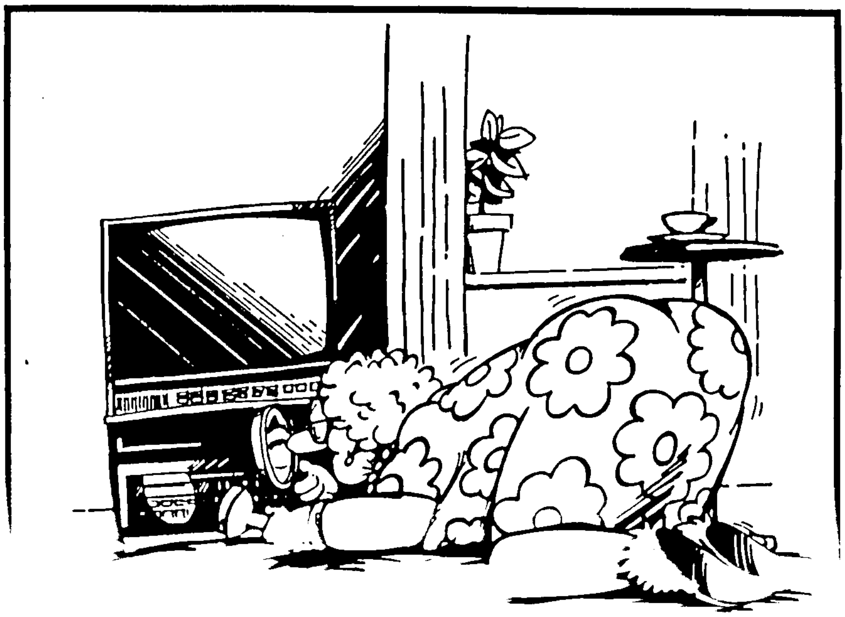


Figure 7 A not-so-funny situation for many users of modern electronic consumer equipment. It is disheartening to see that so few producers actually take the needs of the users into account when designing new consumer equipment. The domestic video-recorder (VCR) has been a particularly troublesome product (Cartoon by Martin Holmes; from [15].)

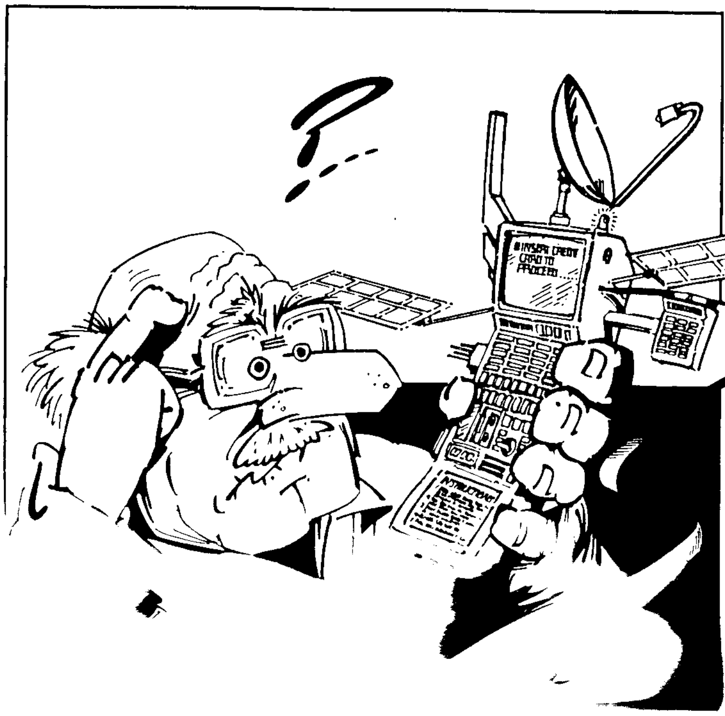


Figure 8 'Over-featuring' is a vice often indulged in by the producers of consumer equipment and it invariably leads to 'under-use': Most of all those clever functions that are invariably crammed into modern communication products are seldom, if ever, used. A bad case of 'creeping featurism'(Cartoon by Martin Holmes; from [15].)

whether over PSTN or ISDN [13,14]. They also demand one-stop-shopping, i.e. to deal with only one company to deliver all that is necessary for a service in one bundle rather than having to deal with a number of different vendors to obtain line access, modem, terminal, communication software, the right connectors and all the other bits and pieces necessary to get a service up and running. The user simply wants to dial a subscriber number to get in touch with the desired party, to type out an e-mail address, to send an e-mail or a web address, to reach a particular home page. Usability will, therefore, become the most important competitive issue after prices have come down. This trend has already been recognised by the larger and more foresighted operators, but other operators are quickly following suit to exploit the commercial advantages offered by improved usability.

If subscribers are to choose a particular service provider it is important that the operator provides services and products which do not require the subscriber to have to memorise long non-intuitive procedures, codes or number strings, or to have to read and understand fat operating manuals. Professor Peter Cochrane of BT Labs in the UK has proposed the ultimate user-friendly paradigm for future ICT (Information and Communication Technology, a term coined by Commissioner Bangeman of the CEC): "Three clicks", "One second" and "No manual"; which means that you should be able to access a desired file or service by not more than three mouse clicks (or three menu choices), that no action should take longer than one second before the user gets a response, and that there should be no need to read user manuals or instructions to be able to use a service or function. This may sound rather utopian, but it is a desirable goal to work towards.

Both the technologies and the necessary human factors knowledge are, by and large, already available, but to achieve many of these desirable goals it will be necessary for some cross sector co-operation; such as between telecommunications engineers, psychologists and programmers and others involved in designing new services and products.

A current hot topic is what is called 'Design for all', 'Universal design' or 'Designing for our future selves', implying that when new products and services are designed these must be usable by as many people as possible. To achieve this it is paramount that human factors experts are brought into the design process as early as possible. The importance of paying attention to human factors issues at an early development stage has been stressed many times, but cannot be repeated too often. Neglect of this may lead to unusable products and services and loss of customers and revenue, requiring costly corrective measures to satisfy the expectations and needs of today's users.

Figure 9 The interactive telephone stored voice services can sometimes be a nuisance to some users

More people live longer, in better health and with more money to spend than at any time before in our history. The market is changing fast as a result of this changing demographic age distribution: As a result of improved nutrition, health care and economy, the population is ageing fast, meaning that a larger and larger proportion of the population will be made up of older people. This development is especially marked in the industrialised countries in Europe and North America, but also in the developing countries of the Third World we see the same trend. The average life expectancy is not always a valid indicator of the growth of the ageing population since it does not take account of the increasingly *skewed* age distribution. We are not only talking about the young elderly (i.e. people in their late 50s or 60s, nor about the middle aged elderly (i.e. those in their 70s and early 80s, but we are talking about the older elderly (i.e. those in their late 80s and 90s) who will also be users of modern ICT services. These user groups cannot be neglected any longer without dire consequences for the service providers. Now when the power of the state monopolies have been broken the subscribers do not have to put up with whatever service or product the monopoly offered. The users will now have real choices, and may not want to spend their money on services and equipment they cannot use. Thus, it becomes more important than ever to create services and products adapted to the special abilities and needs of the various elderly users [2].

The cost of user friendliness

Good user-friendliness does not come free and the cost of creating a good user interface should not be underestimated. Ignoring this issue may have dire consequences. The fixed costs lay in equipping and running a usability lab that in its simplest form may consist of video recording equipment, a computer and various software packages. The cost of this lab, which does not have to be very fancy or expensive, is spread over the various development projects. The variable costs lay mainly in applying the method(s) most appropriate for a specific product during the early development phase, e.g. heuristics, 'Wizard of Oz', cognitive walk-through, user involvement, or any combination of these methods.

The cost of applying such methods can now be estimated quite reliably. The benefits gained by increased sales or by more efficient use of a product with a well-designed user interface can also be estimated with some confidence. By subtracting the costs from the gains, the net gain of developing and applying a good user-interface is derived [16,17,18].

The application of human factors in interface design may have been regarded as a form of esoteric and costly charity, advocated by well-meaning, but somewhat over-enthusiastic philanthropists. Today we know that substantial economic gains can be derived from investing in human factors during product development. The figures cited may vary somewhat, but conservative estimates indicate gains of at least two times the investment – not a bad return on your money by any standard. Most trade journals and popular journals now regularly publish test reviews where user interface and general usability of products are assessed. Buying behaviour is increasingly being influenced and dictated by such product reviews – a product that receives a negative test review stands little chance in the marketplace.

Bell – the first human factors expert on telephones!

When Alexander Graham Bell and Thomas Watson in 1876 developed the first useful system for transmitting the human voice over a wire, Bell did this out of a deep dedication to improve com-

NORWEGIAN TELECOMMUNICATIONS MUSEUM

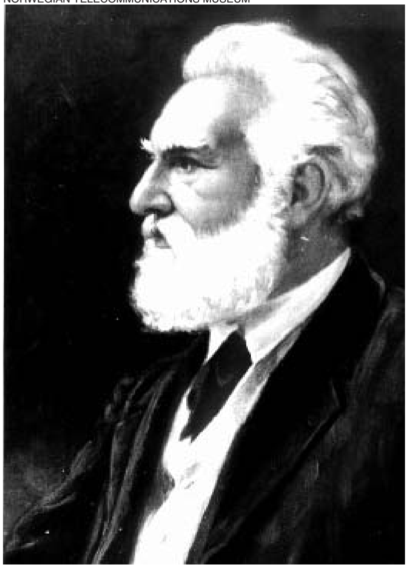


Figure 10 Alexander Graham Bell (1847 – 1922), who together with Thomas Watson invented the first usable telephone in 1876, had also studied the potential users' abilities to operate the telephone system. Knowing well from his work with deaf people how difficult it can be to use technical aids, he wanted the new telephone to be easy to use, thus making him the first 'human factors in telecommunications specialist'

munication for deaf people (see Figure 10). It is, therefore, a deeply tragic irony that no other radical technological innovation has discriminated against deaf people to such an extent as the telephone has. Bell's studies in the physiology of speech and hearing in man provided the basis for his invention, but he did not only study technology and physiology, he also studied the potential users' abilities to operate the new system. This actually makes Bell the first human factors 'expert' in telephony. However, as the technological development gained momentum, the users' requirements soon had to take a back seat. It is important to get the user back in the front seat again, which would also be very much in line with Bell's original intentions.

References

1 ETSI. *Access to telecommunications for people with special needs : Recommendations for improving and adapting telecommunication terminals and services for people with impairments.* ETSI Technical Report ETR 29, 1991.

2 ETSI. *The implications of human ageing on the design of telephone terminals.* ETSI Technical Report ETR 334, 1996.

3 Gagliardi, D. *Report on the audible tones in the telephone service in the EC countries.* Report from contract 48159 DGXIII, European Commission, Brussels, 1993.

4 Anderson, D M, Ferris, A, Mellors, W. Human factors in ETSI : use of tones in public telephone networks. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications, Oslo, 12 – 16 May, 1997.* (ISBN 82-994236-0-0.)

5 ETSI. *European harmonization of network generated tones. Part 1: A review and recommendations.* DTR/HF-01026-1, 1997.

6 ETSI. *European harmonization of network generated tones. Part 2 : Listing and analysis of European, World and standardized tones.* DTR/HF-01026-2, 1997.

7 ETSI. *Minimum man-machine interface (MMI) to public network based supplementary services.* ETSI Telecommunications Standard ETS 300 738, 1997.

8 ETSI. *Guide for usability evaluations of telecommunications systems and services.* ETSI Technical Report ETR 095, 1993.

9 ETSI. *Generic user control procedures for telecommunications services and terminals.* ETSI Technical Report ETR 170, 1995.

10 ITU. *Arrangement of letters, digits and symbols on telephones and other devices that can be used for gaining access to a telephone network.* ITU-T Recommendation E.161 (03/93), 1993.

11 ISO. *Information Technology : keyboard layouts for text and office systems : Part 4 : Numeric Section First Edition.* ISO 9995-4, 1994.

12 ETSI. *Human Factors (HF) : characteristics of telephone keypads and keyboards : Requirements for elderly and disabled people.* ETSI Technical Report ETR 345, 1997.

13 Norman, D A. *The psychology of everyday things.* New York, Basic Books, 1988. (ISBN 0-465-06709-3.)

14 Norman, D A. *The design of everyday things.* New York, Doubleday/ Currency, 1989. (ISBN: 0-385-26774-6.)

15 Collins, S C, Rabbitt, P. *Age and Design for new Technology.* Age & Performance Research Centre, University of Manchester, 1990.

16 Mayhew, D J, Mantei, M. A basic framework for cost-justifying usability engineering. In: R G Bias & D J Mayhew (eds.). *Cost-Justifying Usability.* Boston, Academic Press, 1994. (ISBN 0-12-095810-4.)

17 Harrison, M C, Henneman, R L, Blatt, L A. Design of a Human Factors cost-justifying tool. In: R G Bias & D J Mayhew (eds.). *Cost-Justifying Usability.* Boston, Academic Press, 1994. (ISBN 0-12-095810-4.)

18 Bias, R G, Mayhew, D J (eds.). *Cost-Justifying Usability.* Boston, Academic Press, 1994. (ISBN 0-12-095810-4.)

Knut Nordby is Senior Research Psychologist at Telenor R&D, Kjeller, where he has worked since 1985. He is Chairman of the ETSI Technical Committee Human Factors (TC HF), of which he was a founding member. He is member of the Permanent Steering Committee of Human Factors in Telecommunications (HFT) and Adjunct Professor teaching Man-Machine Interaction at UNIK, Center for Technology at Kjeller, University of Oslo.

e-mail: knut.nordby@fou.telenor.no

Consumers, not Clients – Empowering users of new technology

FRODE VOLDEN

Although new technology has given many individuals with disabilities new opportunities, many of the possible users of such technology do not benefit from these new opportunities. Is this due to poor availability of expertise and technology, or may the hindrance be certain qualities of the service delivery system? Instead of focusing on availability of expertise, there are a number of reasons why we should rather focus on the users' knowledge and skills. Decisions taken by the person himself rather than by others, increase the motivation to succeed in using new technology. A shift of focus from expert-driven decisions to user-driven decisions also has implications for other groups of new technology users, such as elderly people and groups in the society that traditionally have been reluctant to take advantage of new technology. Using the term *consumer* rather than the term *client* may be helpful.

New technology and services, aimed at a growing market of disabled persons, give many individuals new opportunities, and access to information previously unreachable. Estimates indicate that some 10 million people in Europe are potential users of assistive technology due to different impairments. This includes persons with dexterity impairments, visual impairments, hearing impairments, and communication impaired. About 70 % of this population are persons aged 60 and over, see Table 1.

Telecommunication equipment and services as well as other technology have increased many disabled persons' opportunities for independent living. The number of potential users is huge, only a fraction of these are actual users of assistive technology today.

The impact & benefit of assistive technology

A number of studies have shown that many individuals greatly improve several factors associated with quality of life and independence by using assistive technology. As an example of this, a study [2] surveyed 136 individuals with disabilities to evaluate the costs and benefits associated with the use of different kinds of technology-related assistance. The main focus in this report was independence. The individuals were in four age groups and the results indicate a significant impact of AT on many aspects of the respondents' lives, including: the majority of infants with disabilities benefited by having fewer health problems because of AT; nearly 75 % of school age children were able to remain in a regular classroom, and 45 % were able to reduce their use of school-related services; 65 % of working-age persons were able to reduce dependence on family members, 58 % were able to reduce dependence on paid assistance, and 37 % were able to increase earnings. Among elderly persons, 80 % were able to reduce dependence on others, 50 % were able to reduce dependence on paid persons, and 50 % were able to avoid entering a nursing home.

What is assistive technology?

There are a number of definitions of the terms assistive technology and assistive technology services. A simple definition can be: *technology for more independent, productive and enjoyable living*. A more comprehensive definition can be found in the American Individuals with Disabilities Education Act (IDEA).

IDEA defines an *assistive technology* device as:

> *... any item, piece of equipment or product system, whether acquired commercially off the shelf, modified, or customised, that is used to increase, maintain, or improve functional capabilities of individuals with disabilities.* (20 U.S.C. Chapter 33, Section 1401 (25)) [3].

This definition is broad and includes a range of devices from low technology to high technology items as well as software.

Under IDEA the legal definition of *assistive technology* services is:

> *... any service that directly assists an individual with a disability in the selection, acquisition, or use of an assistive technology device.* (20 U.S.C. Chapter 33, Section 1401 (26)) [3].

Assistive technology can be simple or complex. It includes Velcro, adapted clothing and toys, computers, telephony aids, seating systems powered mobility, augmentative communication devices, special switches, assisted listening devices, visual aids, memory prosthetics, and thousands of other commercially available or adapted items. These technology solutions improve an individual's ability to learn, compete, work and interact with family and friends. People use assistive technology to achieve greater independence and to enhance the quality of their lives.

Telecommunication and assistive technology

Telecommunication equipment for the elderly or disabled can be simple additions to existing equipment like enlarged

Table 1 Estimated development of the disabled population in 11 EU Member states (thousands). Source: [1].

	1993, Total	2020, Total	1993, 60 +	2020, 60 +
Dexterity impairments	2,307	2,829	1,645	2,141
Visual impairments	2,189	2,727	1,732	2,254
Hearing impairments	3,419	4,223	2,528	2,290
Communication impaired	1,876	2,308	1,212	1,578
Total	**9,791**	**12,087**	**7,117**	**9,263**

"Big Button"

Big Button is a telephone keypad with extra large recessed keys, and is used in addition to an ordinary telephone. It is designed to assist persons who have difficulties hitting the right key due to spasms or trembling, or to poor eye sight.

It can also be programmed so that it is enough to press one key. In this mode the numbers on the keypad may be replaced by symbols or pictures, making it possible for persons with different cognitive disabilities using telephone themselves.

or adapted keypads for standard telephony, complex solutions for teleworking and "smart-house". Telecommunication services like text-telephony and speech-interfaces rather then keypad, are also examples of telecom based equipment and services for disabled users.

The Internet has given many groups of disabled people access to information they previously had great difficulties accessing. This includes information databases, newspapers, discussion groups, etc. If the user can control and understand a computer using necessary input and output devices, he can in most cases access all the information non-disabled persons can access.

The information-based society, supported by information technology and an advanced communications infrastructure, will probably lead to flexible organisations and new ways of working. An example of this is the rise in the prominence of teleworking. The disabled user who controls a personal computer, will benefit from this development.

The concept "Smart-House" means an adaptation of a domestic environment thanks to computers and telecommunications. The devices in the automated "Smart-House" work together, and many functions in the house is remotely controlled. For instance, doors and windows can open automatically, electrical household appliances can be remotely controlled, and alarms can be set to alert the user or others when certain conditions are satisfied.

Why don't all benefit?

It is pretty clear that assistive technology can improve many individuals' quality of life. What we often see is that persons we thought would benefit from using technology, do not take advantage of these possibilities. Problems with availability and funding as well as available expertise are of course major factors here. But, funding and availability of technology is hardly the problem in our part of the world.

Provision of assistive technology in Norway as well as the other Nordic countries is a key part of our social security system and fully based on governmental funding. Basically, if you document a need for assistive technology, you get it.

Easy access to the technology does not necessarily mean that more people take advantage of the technological possibilities. Several studies have shown a poor relationship between the number of assistive devices provided and the number in actual use [4]. There is reason to believe that there are qualities with the service delivery process itself which explain this lack of relationship.

Expert – client

A focus in (re)habilitation research has been on finding a "perfect match" between user characteristics and assistive technology characteristics. This is of course a necessary part of the service delivery, but probably not sufficient.

Traditionally, the focus in the process of service delivery for disabled persons in need of assistive technology, has been very expert centred. The user has been met with "paternalistic" attitudes, and has been advised what technology he should use, and what to do. Patricia Thornton [5] states that although recent developments in communication technology can reduce dependence on others and facilitate independent living for the disabled, the dominant model of provision emphasises vulnerability, assumes a need for protection and imposes rules about appropriate use of the devices. She argues that this rather leads to disempowerment than empowerment of the user.

Too often the outcome is that people with disabilities have little choice about the services, programs, supports, or accommodations they may receive. Decisions have already been made for them. The empowerment of choice is a key to achieving independence.

Society has changed, and the traditional expert centred service delivery system has of course changed with it. It is considered a democratic right to be informed about decisions concerning yourself. But, this does not necessarily mean that the user is the one in control of the process.

Disabled persons do not have more in common than other people do. They are individuals with their own personality, interests, and priorities. Service delivery systems that do not involve the user and others close to the user will have difficulties in fully adjusting to the user's needs and priorities.

The goal: Independence through empowerment

What we want to achieve when implementing assistive technology is to enable the user to control his or her own life. Independence is the keyword. The popular term empowerment may be useful in this connection. Empowerment is about putting the individual in control, giving him or her the tools to control their own life.

In many countries, most noticeable in the USA, civil rights movements have focused on disabled persons' right to independent living and full integration in society. Taylor, Biklen, and Knoll [6] have tried to describe what independence and community integration for people with disabilities mean. They state that all people belong in community, regardless of their type or degree of disability, and community integration involves participation, community living skills, and advocacy. The individual will be asking questions along with professionals about what it means to be a full participant in the community. This implies that people with disabilities should be put in a position where they run their own programs.

Motivation

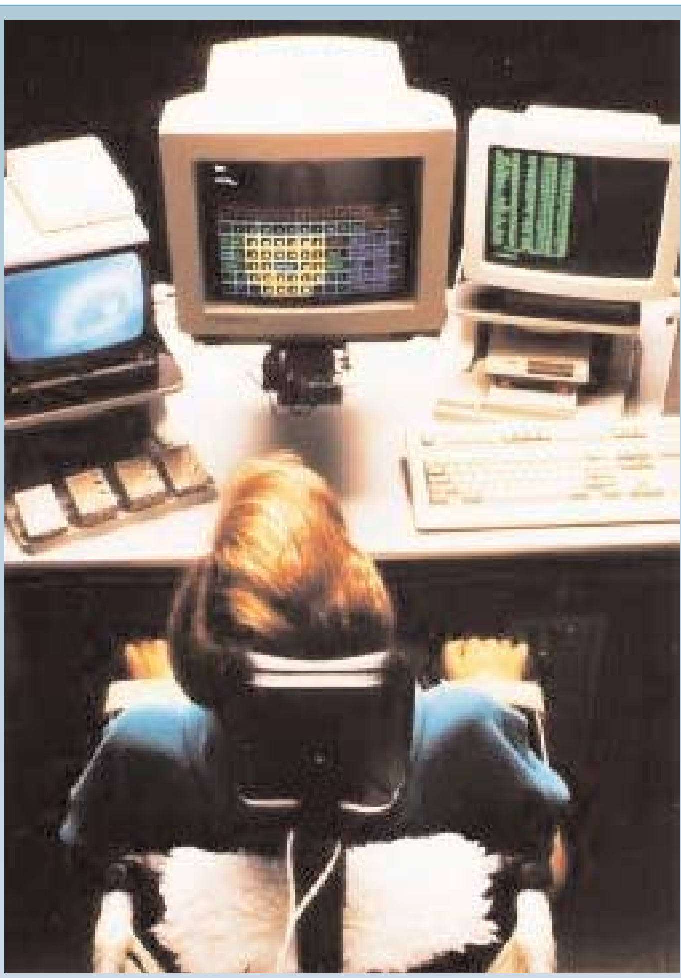
A key issue in the use of new technology is motivation, and an effective way of increasing motivation to take advantage of new technology is created through actual participation from the user in the decision process.

A good match between the user's cognitive and physical characteristics, and technology is necessary, but not sufficient. Starting to use assistive technology is often difficult. In an introduction phase, the assistive devices may decrease rather than increase the user's function. Learning to take advantage of the devices often requires lots of effort from the user and from others near the user. If the selection of technical solutions was just an advice from an expert, it is easy to disregard the expert's competence and stop using the assistive device. If the user himself selects the device, he probably is willing to invest much more effort into learning to use the device and adopt usage of the device in everyday life.

Consumers, not clients

What we would like to achieve is putting the user in charge of his own service delivery process. One step on this road could be changing experts attitudes towards our clients. Using the term consumer rather than client may be one way of achieving this.

Using eye tracking to control appliances



The Eyegaze communication system

The Eyegaze system makes it possible for people with physical disabilities to operate a computer with their eyes. By looking at control keys displayed on a computer monitor, a person can synthesize speech, control his environment (lights, appliances, etc.), type, operate a telephone, run computer software, and access the Internet.

Who uses the Eyegaze system?

Eyegaze systems are used around the United States, Canada, South and Central America, and Europe. Their users are adults and children with cerebral palsy, spinal cord injuries, head injuries, ALS, multiple sclerosis, muscular dystrophy, brainstem strokes, and Werdnig-Hoffman syndrome. Eyegaze systems are being used in homes, offices, schools, hospitals, and long-term care facilities.

How does the Eyegaze system work?

As a user sits in front of the Eyegaze monitor, a specialized video camera mounted below the computer monitor observes one of the user's eyes. Sophisticated image-processing software in the computer continually analyzes the video image of the eye and determines where on the computer screen the user is looking. Nothing is attached to the user's head or body.

The Eyegaze system, developed and sold by LC Technologies, is an example of a "High-end" product that integrates a number of functions that may be useful for a disabled user, for independent living. This is a complex system that requires effort both to learn and to maintain.

More information on this system can be found at their Internet site:
http://www.lctinc.com

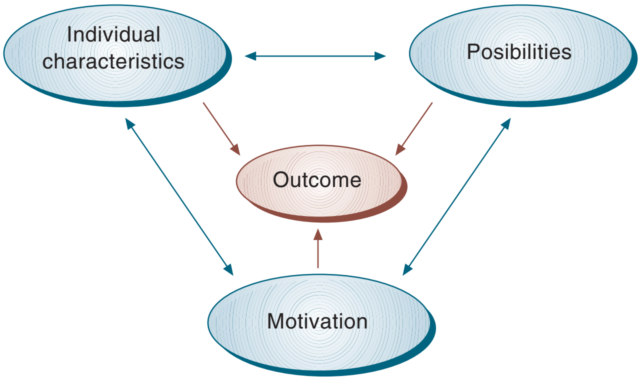
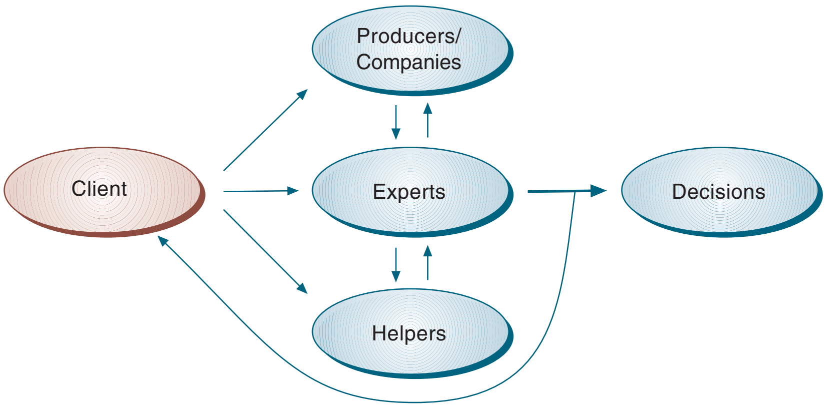
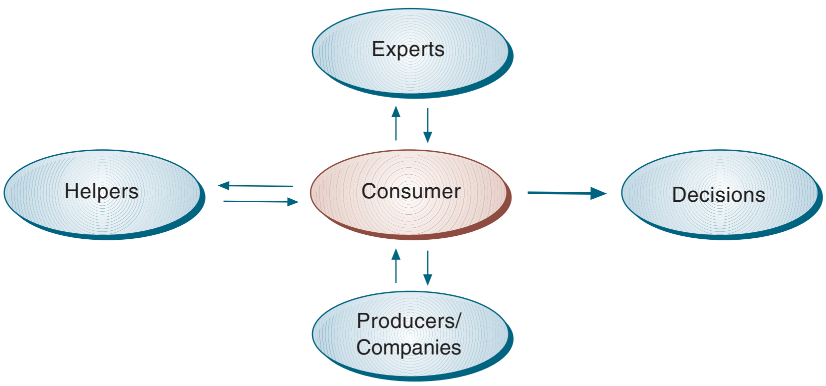


Figure 1 Elements influencing outcome of a service delivery process of assistive technology



a)



b)

Figure 2 The user should be in control of the service delivery process, rather than leaving the decisions to an expert

Nosek [7], in providing support for the word "consumer", says that rehabilitation has a paternalistic past and goes on to suggest that rehabilitation is in the midst of a revolution toward a more humane and creative discipline that focuses on both independence and productivity. She suggests that use of the word consumer symbolises that progress and maintains that use of the word will destroy status quo. Consumer empowerment may be the primary focus of the future.

Older members of society facing adjustments to ageing require services extending beyond traditional views of rehabilitation. Merging the needs of the ageing population with those of disabled persons as advocated by Zola [8] adds strength to consumers' roles. Universal policies recognising that the entire population is at risk for chronic illness and disability will prevent perpetuation of segregated, unequal parts of society. Recognition that everyone with a disability will age, and everyone who ages will acquire one or more disabilities is yet another challenge which broadens the view of those persons providing and receiving services.

Betty Friedan [9] captures the evolving spirit of the ageing in her recent book entitled 'The Fountain of Age'. The older age group intends to have a voice regarding public policies and speak loudly on . issues that concern them. Life beyond fifty no longer means ill health and isolation. Older individuals receiving future rehabilitation services will not be, nor do they view themselves as, a minority or disadvantaged group. The present day strength of such groups as the American Association of Retired People (AARP) symbolises the power of this ever growing population segment.

Access to information

Consumerism requires information; thus, increased access to information accelerates the trend toward rehabilitation consumerism [10]. The communication and technological revolutions have made this a simpler task. Persons with disabilities can access information independently using computers and telecommunications providing access to otherwise inaccessible information [11].

As technological advances enhance opportunities for disabled people, rehabilitation counselling needs to keep pace in effectively matching consumers to technological aids. These advances have brought to the forefront the need for professionals to participate more extensively in rehabilitation assessment and training [12].

Reed *et al.* [13] are advocating what they call a consumer driven model for service delivery, and describe experiences from projects that have tried to focus their service delivery this way. What they call the "local resource team model", is a model they see as well suited for a consumer driven approach.

References

1 Carruthers, S, Humphreys, A, Sandhu, J. The market for R.T. in Europe : a demographic study of need. In: B Ballabio et al. (eds). *Rehabilitation Technology.* Amsterdam, IOS Press, 1993.

2 National Council on Disability. *Study on the financing assistive technology devices and services for individuals with disabilities : a report to the President and Congress,* 1993.

3 The Individuals with Disabilities Education Act (IDEA), 1997.

4 Wormnæs, S. *En større verden : datateknologi for personer med bevegelses- og talevansker.* Oslo, Universitetsforlaget, 1992.

5 Thornton, P. Communication technology : empowerment or disempowerment. *Disability, Handicap and Society,* 8, 4 Dec, 339–349, 1993.

6 Taylor, S J, Biklen, D, Knoll, J (eds). *Community integration for people with severe disabilities.* NY, Teachers College Press, 1987.

7 Nosek, M A. A response to Kenneth R Thomas' commentary : some observations on the use of the word "consumer". *Journal of Rehabilitation,* Apr/May/Jun, 9–10, 1993.

8 Zola, I K. Ageing and disability : toward a unified agenda. *Journal of Rehabilitation,* Oct/Nov/Dec, 6–8, 1989.

9 Friedan, B. *The fountain of age.* New York, Simon & Schuster, 1993.

10 Inlander, C B. Trends in medical consumerism. In: A M Rees (ed). *Managing consumer health information services.* Phoenix, AZ, Oryx Press, 3–14, 1991.

11 Fullmer, S, Majumder, R K. Increased access and use of disability related information for consumers. *Journal of Rehabilitation,* Jul/Aug/Sep, 17–2, 1991.

12 Hale, P N Jr, Shipp, M. Personal transportation : a challenge to the rehabilitation community. *Journal of Rehabilitation,* Apr/May/Jun, 8–9, 1989.

13 Reed, B J, Fried, J H, Rhoades, B J. Empowerment and Assistive technology : the local resource team model. *Journal of Rehabilitation,* 61, (2), 30–35, 1995.

Frode Volden is Research Psychologist at Telenor Research & Development, Kjeller

e-mail address: frode.volden@fou.telenor.no

Speech quality, usability and service quality

NORMAN GLEISS

1 The concept of speech transmission quality

The perceived sound quality of speech transmitted and reproduced over a technical system, in particular the telephone system, may be described by a number of attributes, which ultimately affect the usability of telephony services. These attributes may be derived from technical, physiological or psychological considerations and be based on empirical data from these domains.

1.1 Information theory

Technically, some kind of correlation may be expected between speech quality and the ratio of the amount of information transmitted over a specific system to the total information content of the speech spectrum. If this ratio is <1 it is reasonable to believe that speech quality is degraded; at least that should be true for analogue signal transmission, disregarding the redundancy of speech information and the possibility of intelligent coding techniques which might reduce the redundancy in a way that is not noticeable to the listener.

According to Shannon's formula, the maximum amount of information, C, which can be transmitted over a channel, is proportional to the bandwidth B, the speech level S, and, inversely, to the noise level N:

$$C = B \cdot \log_2 (1 + S / N) \text{ bit/s}$$

There is experimental evidence that speech quality Q is indeed a function of the speech and noise levels and of the bandwidth, i.e. $Q = f(S, N, B)$. The possible relation to the transmission of information has been investigated by Richards and Swaffield [1]. Suppose that the information H_i bit/s is fed into the sending end of a speech transmission link having a capacity of C bit/s, and that the amount H_o reaches the receiving end each second. With insufficient channel capacity, i.e. $C < H_i$, we get $H_o / H_i < 1$, but when $C > H_i$ we have $H_o = H_i$. A high-quality circuit could have a certain excess capacity $C - H_i$. With these assumptions, the ratio of H_o / H_i as a function of channel capacity would follow the curves shown in Figure 1. Here the linear theoretical function presumes a perfect coding and decoding of the information stream. The practical performance will follow the dashed curve because of the less than ideal electro-acoustical coding in the telephone set.

However, the channel capacity required for speech transmission is actually determined by the rate at which the hearing sense can resolve the acoustic information received rather than by the information content of the input signal. The maximum receive rate is estimated to be 15 kbit/s. A value of C of this magnitude corresponds to an analogue telephone connection with a bandwidth of B = 300 – 2800 Hz and a signal-to-noise ratio of S/N = 16 dB. These values would, however, be outside the range recommended by ITU-T for international telephone transmission, and the quality of such a connection would be judged as very bad. Actually, a capacity of 50–70 kbit/s is needed for good speech transmission. On the other hand it is now possible to obtain satisfactory speech transmission quality at 16 kbit/s or even less by advanced digital coding techniques.

A different kind of limitation is the rate at which the brain can handle the information coming from the ears. This is very much smaller than the hearing resolution: only of the order of 10 bit/s. It should therefore be possible to transmit at least the semantic information contained in a speech signal over a channel with a considerably lower bit rate, if a suitable coding procedure could be found. In fact, a telex system having a capacity of 50 bit/s is capable of transmitting the semantic content of a message in the form of text at about the same speed as the words are spoken. (Speech in a language composed of $2^5 = 32$ phonemes (5 bits) and spoken at a speed of 10 phonemes per second implies a transmission rate of 50 bit/s.)

1.2 The hearing sense

The hearing sense has two characteristics which are of particular importance for the perception of speech over a telephone system. One is the sensitivity in relation to frequency, which can be described by "isophones" presenting the sound pressure level required to give a specific loudness sensation as a function of frequency. The threshold of hearing is the lowest curve of the isophone family, corresponding to zero loudness. The threshold moves upwards in level when audible noise is added and becomes then a masked threshold.

The other important characteristic is the loudness growth function, which presents the perceived loudness of a sound as a function of sound pressure level. This function has a shape that to some degree depends on the spectrum of

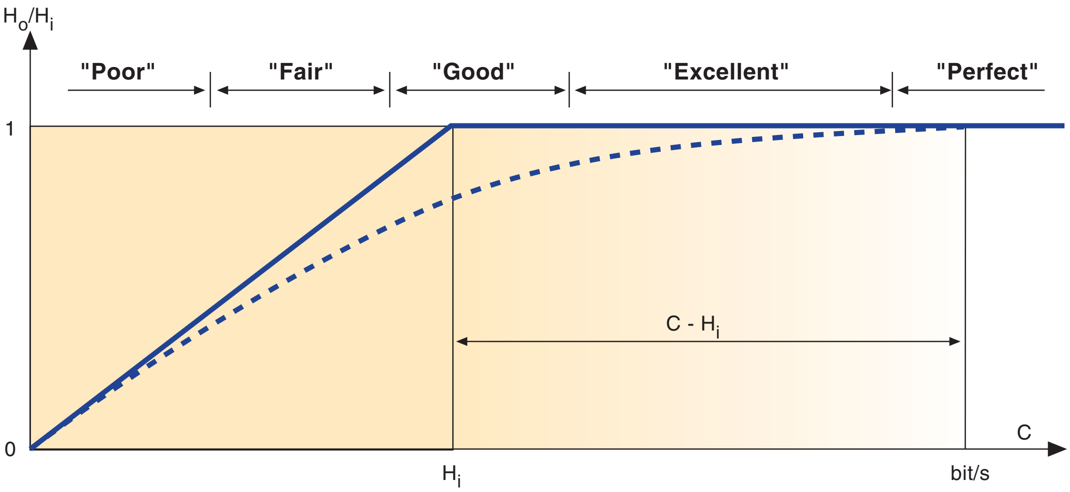


Figure 1 Performance of a speech transmission link as a function of channel capacity. Input information rate = H_i; output information rate = H_o. Full line: theoretical function; dashed line: practical performance

the sound. The loudness growth function for speech is a special case which will be discussed later in this paper.

With consideration to these characteristics of the hearing sense, the useful part of the information in transmitted speech may be expected to be that part of the speech spectrum which is left over after being affected by the loss, the noise and the band limitation introduced by the transmission system. Figure 2 shows this part as the area of information content in a level/frequency diagram, where the lower contour is defined by the hearing threshold of the listener and includes the masking effect of supra-threshold noise.

From this kind of presentation it is possible to calculate the intelligibility of the received speech in dependence of the quantities *B, S* and *N*. The calculation of Articulation Index is the best known method for this purpose. Here the useful spectrum is divided into a number of frequency bands that each contribute to articulation or intelligibility by a certain weighting coefficient that can be determined empirically. This method was originally proposed by Collard (1930) and has later been developed by, among others, French and Steinberg (1947).

This acoustical description, involving the physiological characteristics of speech production and hearing, is also used in recent models for calculating not only intelligibility but also the overall perceived quality of speech impaired by a transmission system. These models therefore include empirical data from subjective tests where the quality is rated on scales representing different perceptual dimensions. There are furthermore specific models which are used to calculate only the loudness of speech. The results may be presented as "Loudness Ratings". In this case the non-linear function relating the growth of loudness with speech level is incorporated in the model.

An overview of calculation models by which perceived dimensions of speech can be calculated from electro-acoustic data, basically measurements of frequency response (bandwidth) and speech and noise levels, is given in ITU-T Supplement 3 to the P-series Recommendations [11]. Further work in the development of calculation models for speech quality, including also the effects of digital coding, is presently going on in ITU-T/Study Group 12. One recent result is Recommendation G.113 [9].

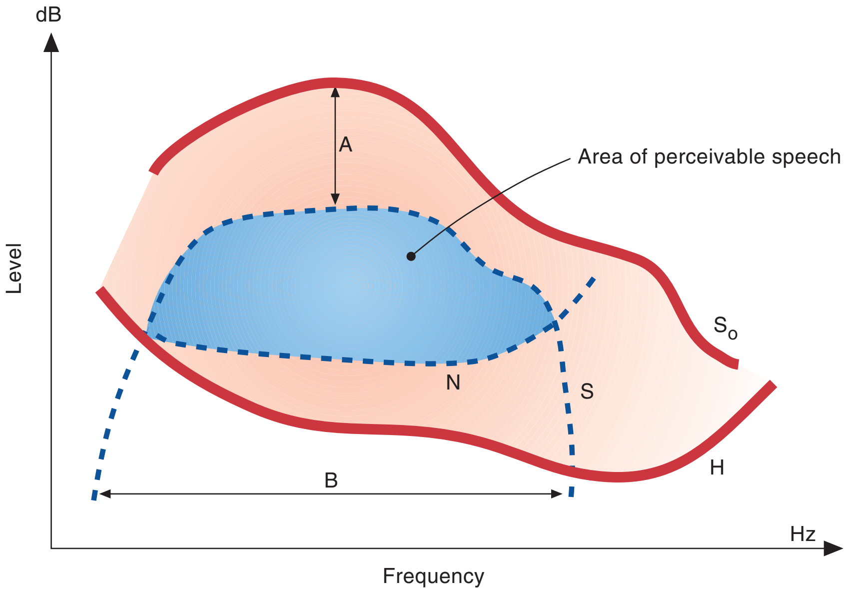


Figure 2 Principles of the calculation of intelligibility from the received speech spectrum (S), hearing threshold (H), and noise level (N). The area of perceivable speech has the received speech spectrum (the sum of the input speech spectrum, S_0 and the frequency-dependent attenuation in the transmission channel, A) as the upper boundary and the masked hearing threshold (the sum of the threshold and the noise level) as the lower boundary

1.3 Perceptual dimensions

A large number of experiments have been carried out over the years on analogue telephone transmission systems in order to investigate the relation between physical system characteristics, basically *B, S* and *N*, and perceived quality. In general, these experiments have been conversation or listening tests where the system parameters were varied in all dimensions, and subjects were asked to judge the overall quality on an opinion scale. By different methods of statistical analysis, in particular factor analysis, it has been possible to identify the most important factors which make up the general concept of speech quality.

Not surprisingly, the dominant factor is *Loudness*. In fact, most of the engineering effort during the first 50 years of the

Table 1 Relation between the perceived quality of speech and the physical system characteristics

Physical factors	Perceptual factors			
	Overall quality	Loudness	Intelligibility	Naturalness
Speech level	+++	+++	+++	+
Noise level	+++	+	+++	+
Bandwidth	+++	++	++	+++
Distortion	++		+	++
Sidetone	+	+ *(Talker sidetone)*	+ *(Listener sidetone)*	+
Delay	++			+
Echo	++		+	+

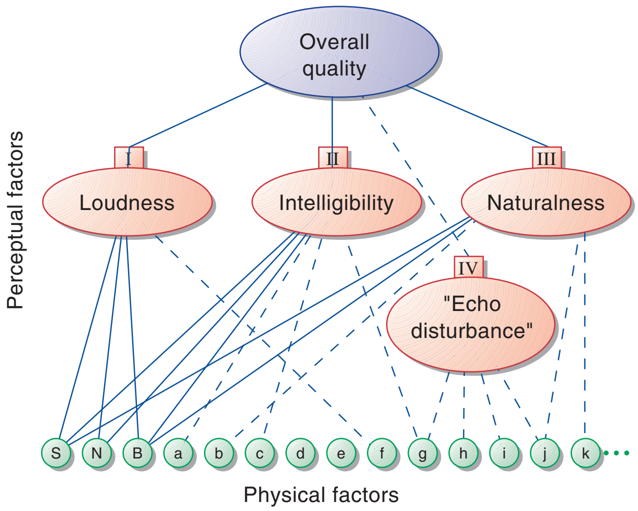


Figure 3 Structural model of perceived speech quality and its components

history of telephony has been devoted to produce a sufficiently high listening level at the receiving end of the telephone connections. The second factor in importance is *Intelligibility*. Again this would be expected, because the principal objective has been to transmit intelligible speech. Only as number three comes a factor which may be called *"Naturalness"*. This factor has to do with the non-semantic content of speech and involves personal characteristics such as the identity and the state of mind of the speaker.

The basic physical factors are, as expected, speech level, noise level and frequency characteristics or bandwidth. Non-linear distortion, sidetone, delay and echo effects can be considered as secondary factors which contribute to a minor degree to the overall quality. The relation between the physical and perceptual factors may be represented as in Table 1. The strength of the individual relation between factors is approximately as indicated by the number of crosses. These relations are discussed in more detail in the next section.

The structural relationship between the perceptual and physical factors is illustrated in Figure 3. Apart from the basic factors *S, N* and *B*, which all have a specific relation to the three most important perceptual dimensions, there may be a large number of other factors which either affect the same dimensions or only one specific additional dimension. For example, talker echo means that a speaker hears his own voice again after a certain delay and with a certain attenuation (the "echo loss"). The perceived disturbance of an echo is a complex function of echo loss and delay and may form its own quality component, designated by *IV* in the figure, because it may not significantly affect either loudness, intelligibility or naturalness.

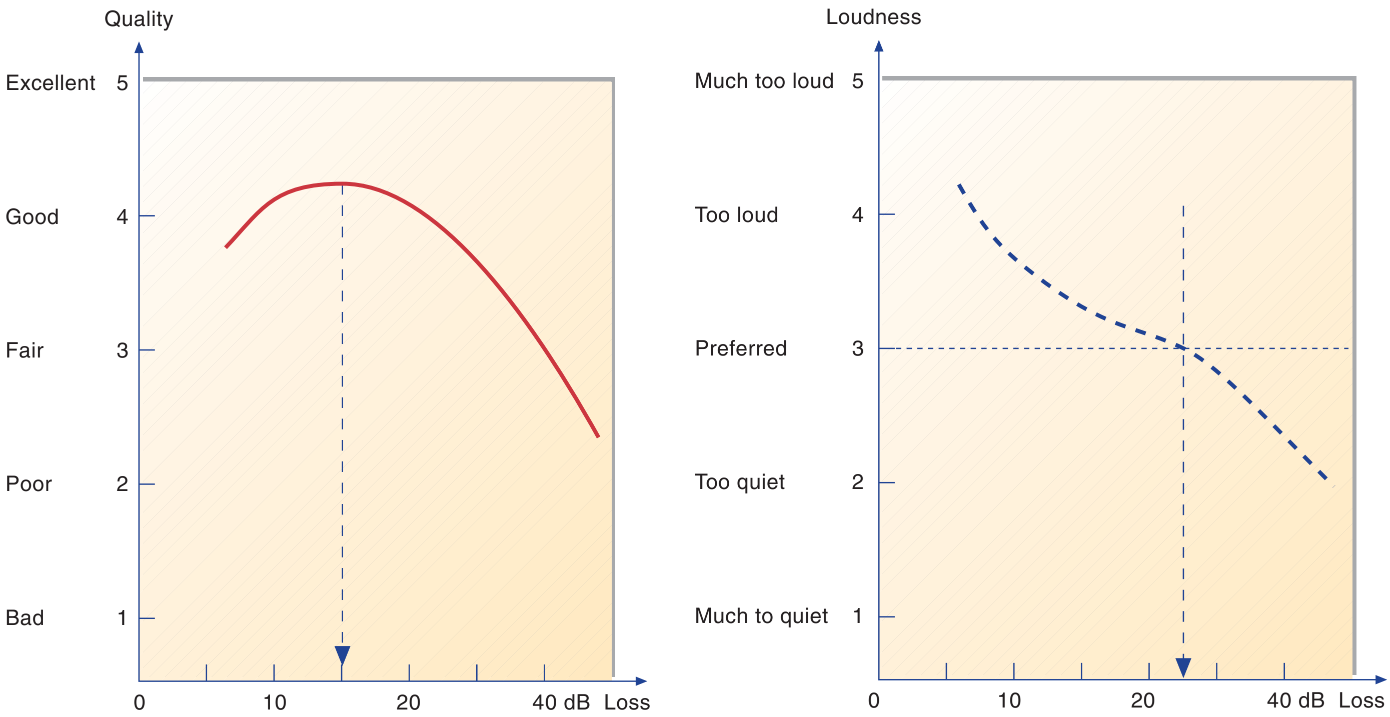


Figure 4 a) Speech quality; b) Loudness, as functions of transmission loss (decreasing speech level); general principle

2 Quality dimensions

2.1 Loudness of speech

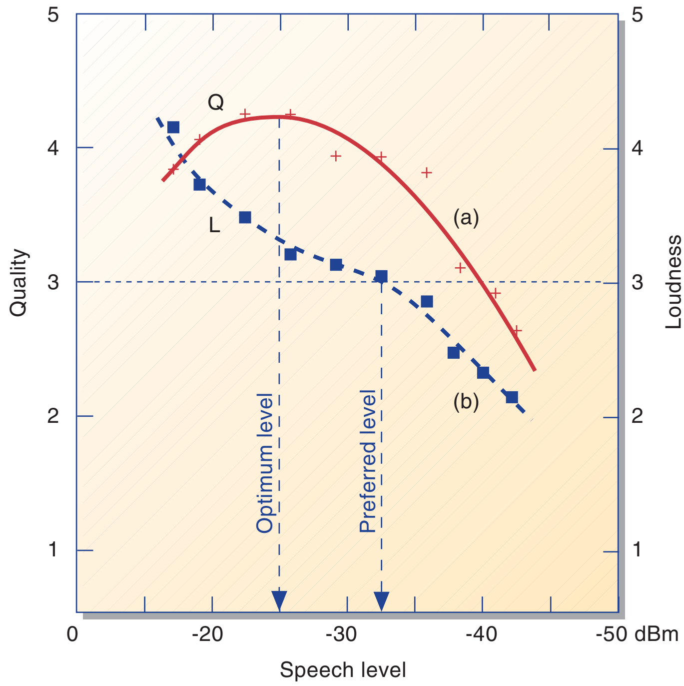


Figure 5 a) Speech quality; b) Loudness, as functions of decreasing speech voltage level, as obtained in a listening test with 20 subjects over a modern telephone set

Loudness is generally the dominating factor for the perception of speech quality. A typical curve relating perceived quality to loss in a telephone transmission system is shown in Figure 4a (increasing loss means decreasing sound level at the listener's ear). The sound quality has a maximum at a specific sound level (corresponding to a certain loss, or "Loudness Rating", in the system). The perceived quality falls off at both sides of the maximum, although steeper in the direction of increasing loss, coming down to zero when the speech is no longer audible.

At the two sides of the maximum, points of equal quality may be found; obviously one is too loud and one is too weak, and the optimum is somewhere in between. For further analysis, the Quality function may be compared with a similar function presenting loudness against loss (Figure 4b). It might be expected that the *optimum* speech level (at "optimum" loss) coincides with the most comfortable or preferred listening level, and in fact in earlier literature the diagrams are often drawn in that way (e.g. [3]). However, the results of tests where subjects are asked to rate quality and loudness independently in a modern telephone set, may look as in Figure 5. Similar results, contributed to Study Group 12, have been obtained in different laboratories. Therefore, the curves in Figure 4 demonstrate a typical effect, implying that the optimum speech level is significantly higher than the *preferred* listening level. Differences in the range of 3–8 dB have been found, depending, among other things, on the type of telephone. This effect may be interpreted to mean that maximum speech quality is perceived to occur at a higher level than the most comfortable listening level. Similar results have been reported for users of hearing aids.

The global telephone network is designed so that as many of the users as possible get a listening level that is close to the optimum level. With national standard telephone sets complying with ITU-T Recommendations the sound pressure level will then be in the range of 80–85 dB. This value is considerably higher than the average listening level, which is about 65–70 dB, during a conversation between two persons at a convenient distance in an ordinary room (with a certain noise level). The same level of 65–70 dB is considered as a comfortable level when listening to a loudspeaker telephone. One reason for this difference of about 15 dB is that the preferred listening level increases with room noise, but the greater part can be derived from the difference between monaural and binaural listening, see Figure 6. In the figure, the estimated loudness for continuous speech is presented on a logarithmic scale as a function of the sound pressure level at the ear for one and two ear listening, as measured by Gleiss and Goldstein [4]. It can be seen that in order to produce the same loudness for monaural listening as for binaural listening, the level has to be raised by about 12 dB. Therefore, it is very likely that the above mentioned difference of 15 dB can mainly be attributed to the monaural/binaural difference. On the other hand, going from one to two ear listening with the same level at the ear appears to increase the loudness only by 1.7 : 1, while earlier estimates have been 2 : 1.

The shape of the curves in Figure 6 resembles the loudness curve in Figure 4b. The reduced slope in the middle region compared to the ends appears to be typical for speech and can be explained as a psychological "constancy effect": a change of sound pressure in the range where conversational speech is normally presented to the ear is to some extent perceived as a change in distance to the talker rather than in loudness. The perceptual image is similar to that of a talker changing the spatial distance to a listener in the same room while maintaining his speaking level; in a real conversation situation the listener would correctly interpret the change of sound level at his ear as a change in distance and not in loudness, whereas in a telephone conversation the same interpretation becomes an illusion. (For earlier generations of telephone users it was perhaps reality, because, in the analogue network, loss generally increased with the geographical distance between the partners in a conversation and so contributed to the sensation of physical distance.) This effect enhances the importance of being in the vicinity of the optimum listening level in order to give an impression of social presence to the partners in a telephone conversation.

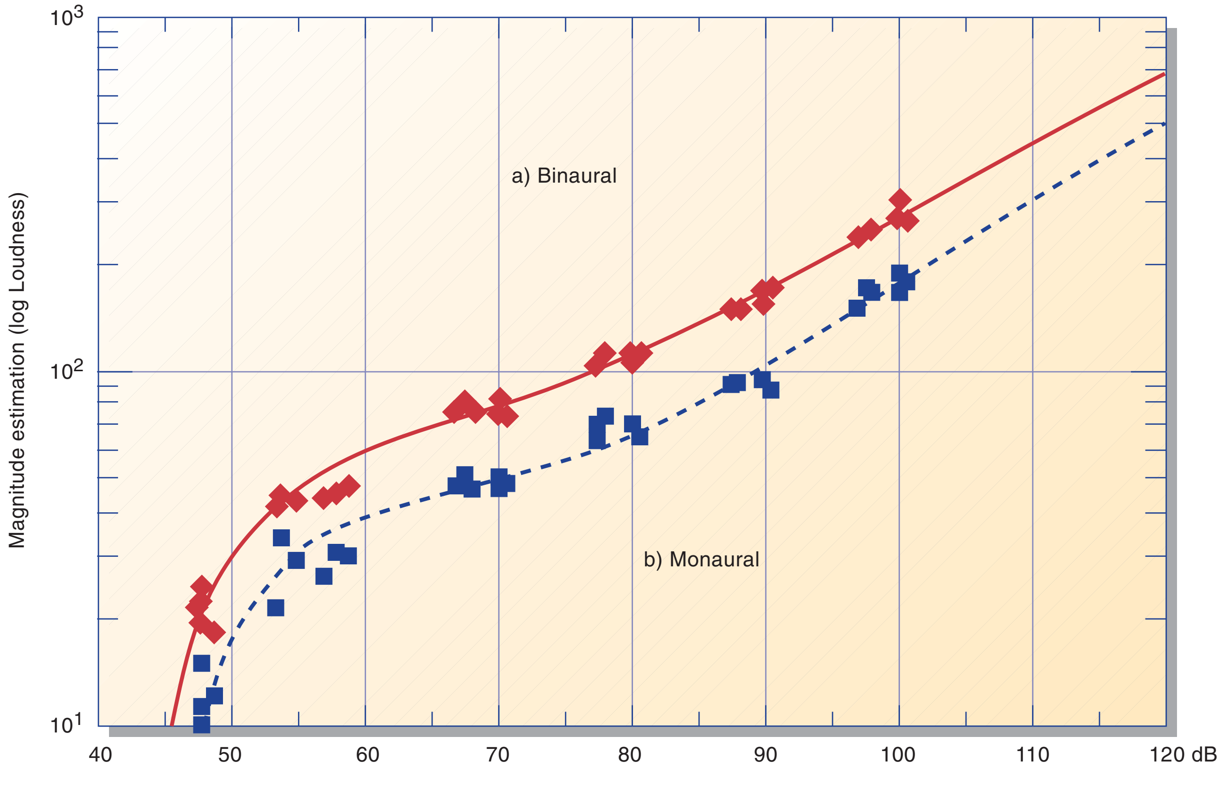


Figure 6 Loudness as function of speech level at the ear, for a) two ear, and b) one ear listening (curve fitted to data from Gleiss and Goldstein)

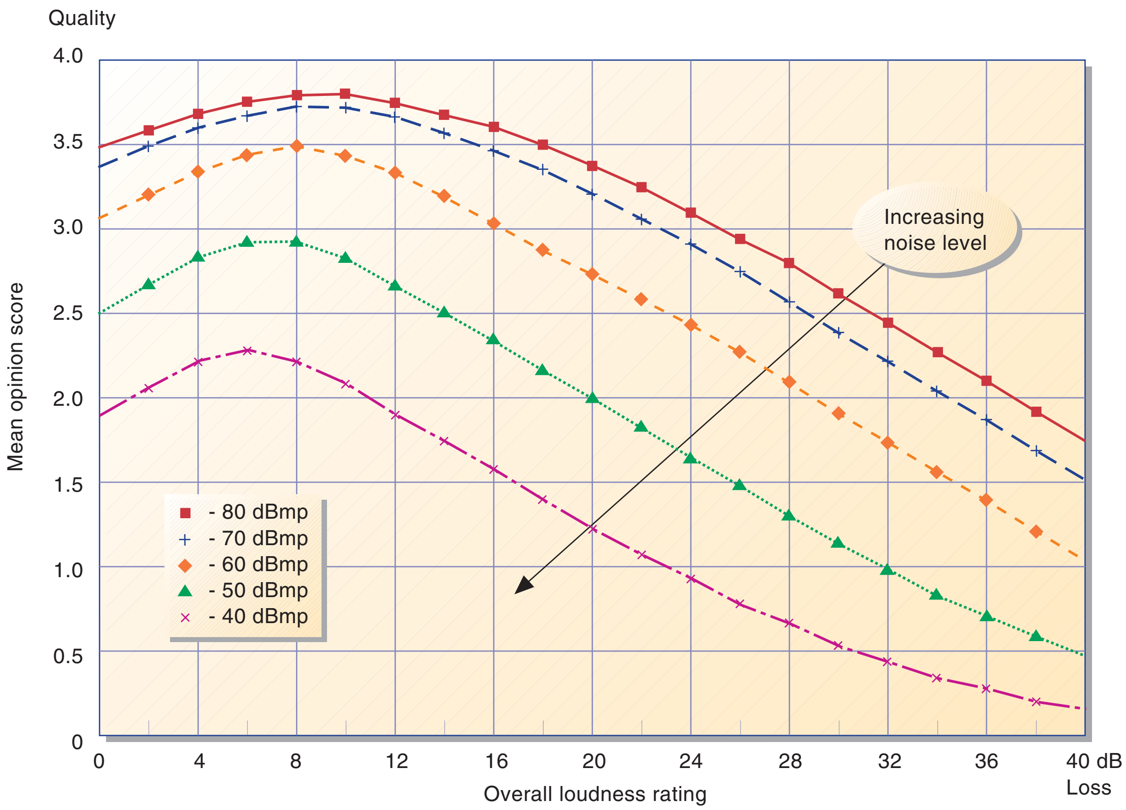


Figure 7 Perceived quality, as mean opinion scores on an overall quality scale, as a function of transmission loss (decreasing speech level) with noise level as a parameter

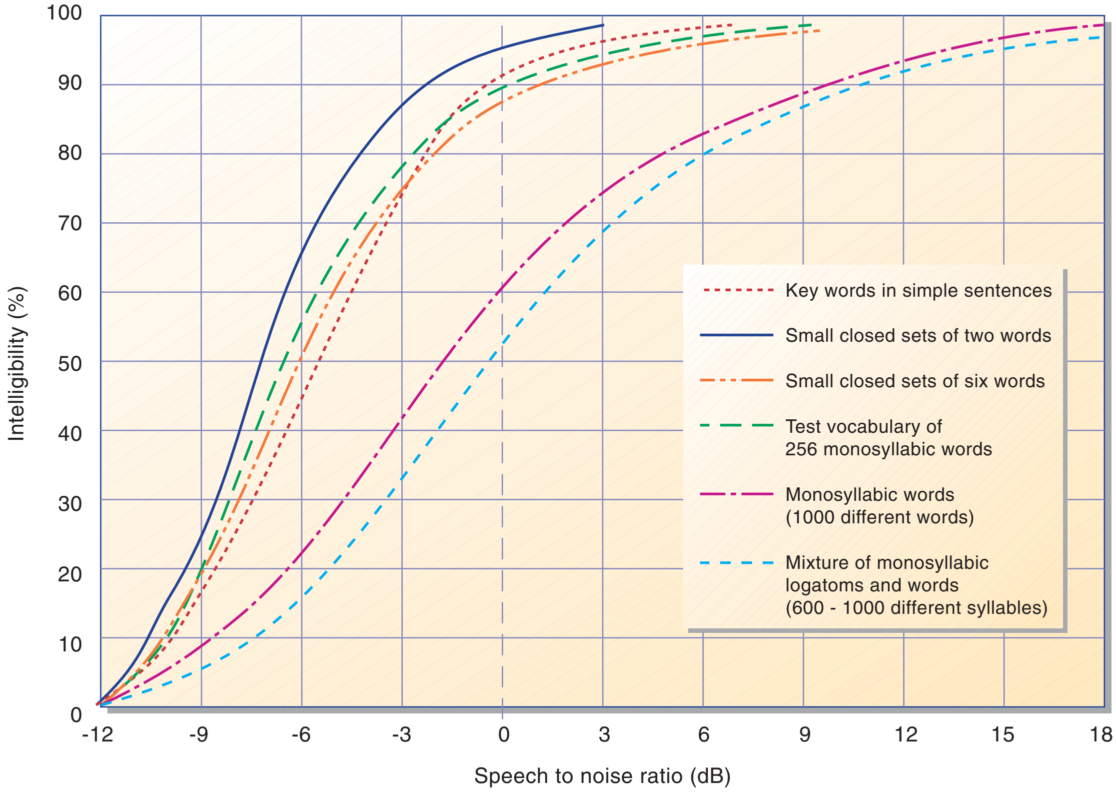


Figure 8 Intelligibility as a function of speech-to-noise ratio (S/N) for various types of speech material (From ISO)

Another effect to consider in the planning of telephone systems is the influence of noise. Line noise as well as room noise can reduce the signal-to-noise ratio for speech at the listener's ear. Consequently, intelligibility may also be reduced, and therefore the optimum listening level in the presence of noise may be shifted towards higher speech levels. This effect is illustrated in Figure 7, showing a family of curves with different noise levels as parameters. The uppermost curve without noise is in principle the same as in Figure 4a. It is clearly seen that for higher noise levels, the optimum points are shifted to the left, i.e. to higher listening levels (lower loss values). As a consequence, "isopreference" contours can be found which contain points on different curves having the same perceived quality (horizontal lines in this diagram). At least over a limited range (on one side of the optimum and for higher noise levels), these isopreference contours will correspond to constant *S/N* values.

2.2 Intelligibility

Intelligibility, or articulation[1], is generally measured by counting the proportion of correctly understood sounds, syllables, words, or sentences in a listening test with a group of subjects. In principle, the results for a system without bandwidth limitation or other distortion will come out as in Figure 8. The curves describing intelligibility as a function of *S* or *S/N* have a sigmoid shape going from 0 % and gradually approaching 100 %. They are centred around the 50 % level, which may be called the threshold of intelligibility. These general results mean that intelligibility primarily depends on the signal-to-noise ratio but also on the speech material itself. Intelligibility tests are in fact performance tests, which measure the performance of the listener as well as of the transmission system. The expectations of the listener and the information content of the speech material presented in relation to the listener's own vocabulary have a strong influence on the measured intelligibility. However, with this in mind it is possible to at least rank order different systems in quality with a given listener group and a specific type of speech material. On the other hand, in order to come to conclusions on an absolute basis, the average performance of a large number of listeners has to be assessed.

Formal articulation tests are rarely used nowadays, not only because they are time consuming but even more because transmission systems in general are required to be well outside the saturation level of 100 % of the curves in Figure 8, in fact outside the range covered in the diagram. In practice, telephone systems should have an intelligibility of at least 90 % for syllables and almost 100 % for sentences. It is therefore not possible to differentiate between actual systems by articulation tests. When new systems or components,

[1] *In technical language, articulation usually refers to identifying sounds and nonsense syllables, while intelligibility refers to understanding words and sentences.*

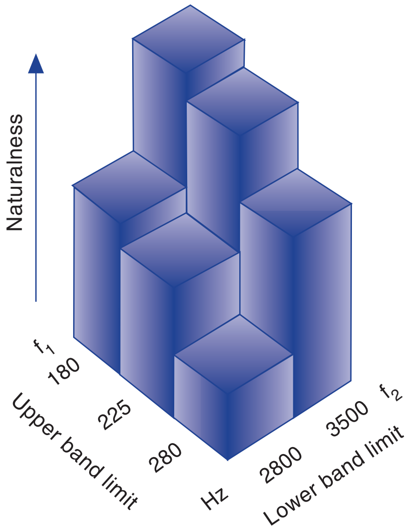


Figure 9 Perceived naturalness of speech transmitted over frequency bands with different upper (f_1) and lower (f_2) cut-off frequencies. Results from paired comparison listening tests; male and female voices

e.g. speech coders, are introduced it is generally sufficient to rate "intelligibility" or "listening effort" directly. This can be done with surprising accuracy and with much less effort than articulation tests.

Suitable rating scales for this purpose may have a zero point labelled *"No meaning understood with any feasible effort"*, while the upper end of the scales may be labelled *"Complete intelligibility"* or *"Complete relaxation possible; no effort required"*.

Intelligibility is also dependent on the frequency response or bandwidth of a transmission system. The conventional telephone band is limited to 300 – 3400 Hz. Historically, the band limits were chosen as the frequency limits for low-pass and high-pass filters which gave the same reduction in intelligibility. Extending the band limit upwards in frequency does not dramatically improve intelligibility. An extension up to 7000 Hz, which is the standardised upper limit for ISDN broadband telephony, certainly increases the intelligibility for some consonants, especially fricatives like *f* and *s*, but at the same time the relative improvement in overall sound quality is much more noticeable. This improvement in quality is even more pronounced when the frequency band is extended downwards. It appears that changes in bandwidth, at least outside the conventional telephone band, affect the factor of naturalness much more than intelligibility. Naturalness should therefore be treated as a separate issue, with equal importance for speech quality as intelligibility.

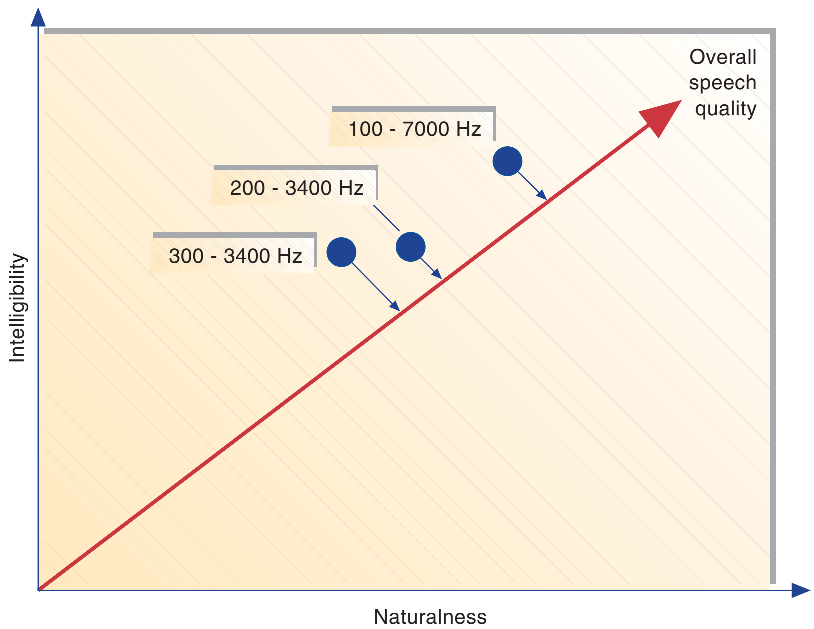


Figure 10 Naturalness and intelligibility ratings for three different speech transmission bands and their projection on a hypothetical quality scale

2.3 Naturalness

It has already been mentioned that the concept of naturalness covers speaker recognition and the timbre of the voice but also other more subtle non-semantic components such as the perception of the speaker's emotional status as reflected in the character of the voice. Recent experiments have shown that facial expressions like smiling or frowning not only affect the acoustic speech spectrum in a measurable way but also that listeners can detect and identify these expressions from the voice character without seeing an image of the speaker. On the other hand, seeing the face of the talker, as in videotelephony, can significantly improve intelligibility, at least under adverse acoustical conditions, and also contributes positively to the perceived speech quality.

The frequency response of a speech transmission system affects intelligibility and naturalness in different ways. Naturalness is in most applications improved by a full bass reproduction and also a certain degree of treble boost, particularly if the upper cut-off frequency of the transmission band is not sufficiently high. Intelligibility, on the other hand, may be improved by a bass reduction and a further emphasis on higher frequencies. In general, naturalness seems to be more dependent on the presence of the fundamental frequency (F_0) and the first formant (F_1) in the reproduced speech spectrum, while intelligibility is more dependent on the higher formants in the voice spectrum. However, these effects depend to a high degree on the speakers' individual voice spectra. In particular, there are significant differences between male and female voices as regards the magnitude of the effect of bandwidth restriction on speech quality.

A detailed analysis of the effect of changes in bandwidth and frequency response based on extensive tests at the former Swedish Telecom Administration was made by Gleiss [2]. It was found that shifting the band limit by 100 Hz at the lower end affects naturalness much more than a similar shift at the upper end. Therefore, for example the frequency band 200–3050 Hz is perceived as clearly superior to the equally wide band 300–3150 Hz. If instead the limits of the frequency band are moved in steps of one third-octave, results are obtained as shown in Figure 9 for six different bands. It appears that the impairment caused by decreasing the upper cut-off frequency from 3580 to 2800 Hz can almost be compensated for by decreasing the lower cut-off frequency from 280 to 380 Hz, i.e. a decrease of 100 Hz at the lower end compensates a decrease of 750 Hz at the upper end with respect to the naturalness of the transmitted speech. A conclusion is that the conventional telephony band of 300–3400 Hz is not the optimum choice and a displacement of the band 100 Hz downwards would improve the speech quality significantly.

Further examples of the effects of specific modifications in the frequency response of a telephone system are given in Figures 10 and 11. Figure 10 shows how three different transmission bands are positioned in a diagram where the coordinates are Intelligibility and Naturalness. It is seen that an extension of the band downwards from 300 to 200 Hz increases naturalness but has no noticeable effect on intelligibility. Increasing the upper band limit from 3400 to 7000 Hz (approximately one octave) makes a larger contribution to intelligibility but also to naturalness.

The rank order in overall speech quality can be seen from the projections on the hypothetical quality dimension, which goes at an oblique angle between the orthogonal dimensions. Splitting up speech quality into its components obviously gives more detailed information about the perceived difference between the systems. (In this kind of test the loudness has been equalised between the transmission bands, so that loudness is not involved as a separate dimension in the evaluation.)

Figure 11 (from [1]) demonstrates how the effect of two different slopes of the frequency response (either flat or with a +6 dB/octave preemphasis) within the band is related to two extreme acoustic conditions (reverberant or anechoic) at the send or receive end of a telephone connection. Again, naturalness and intelligibility have been rated separately. The results indicate that naturalness is always better with a flat response but that intelligibility increases somewhat in the two cases with reverberation at the send end when the preemphasis is applied. The test results also prove that naturalness and intelligibility are indeed orthogonal dimensions.

Other experiments, e.g. by McGee [5], have demonstrated that naturalness and intelligibility (at constant loudness) are the best descriptors for the perceptual dimensions that are the fundamental components of speech quality. McGee used semantic differential scales for evaluating a large number of system conditions. He found that the Naturalness factor was highly correlated with such scales as Big–Small, Wide–Narrow, Full–Thin, Pleasing–Annoying, and Natural–Unnatural. The Intelligibility factor was more related to Bright– Muffled, Sharp–Dull, etc. and, of course, to scores calculated from articulation index models.

The ultimate consequence of these findings is that the quality of speech systems under test can be rated directly on the two dimensions of Naturalness and Intelligibility, for example on 5-grade scales going from *"No Naturalness (or Intelligibility) at all"* to *"Complete Naturalness (or Intelligibility)"*. Figures 10 and 11 are examples of actual applications. By this method, the redundancy of rating scales in a questionnaire, and with that the total test time, can be considerably reduced.

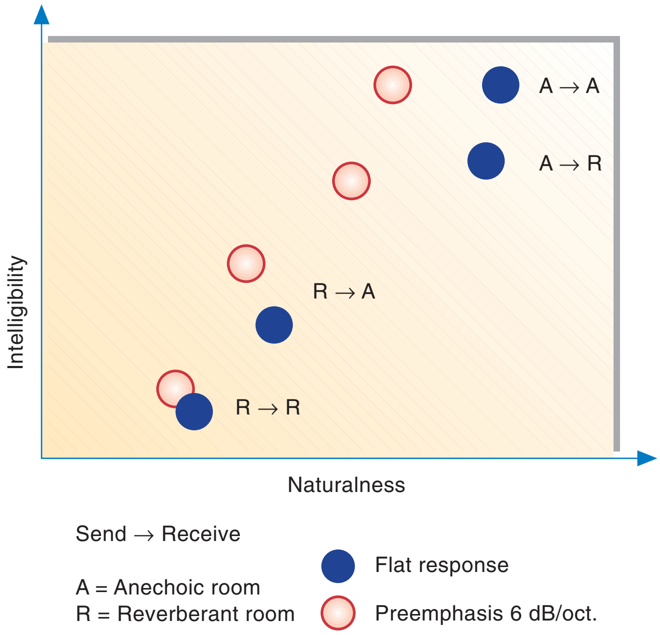


Figure 11 Naturalness and intelligibility ratings for a loudspeaker telephone system in different room acoustics. The send and/or receive ends are in anechoic or reverberant room conditions. The microphone frequency response curve is either flat or has a 3 dB/octave preemphasis

3 Usability

After this overview of the concept of speech quality and its components, it is appropriate to make a comparison with usability concepts and look for possible relations between the two. Many different definitions of usability are found in literature, some of which are language specific and cannot be easily translated into other languages. Furthermore, usability definitions are often operational, i.e. what they mean depends on how they are measured.

In 1974 the Communications Studies Group, CSG, in London evaluated the usability of a few types of conference telephones for a Scandinavian working group. They performed a large number of laboratory tests with different groups of users, who had to use the various telephones for the same kind of tasks. Both the task performance and the attitudes of the users to the equipment were recorded.

In the reports from these experiments, CSG applied the concepts of *Effective-*

ness and *Acceptability* to the two principal dimensions for structuring the results. By effectiveness they meant how well the task was solved by the respective equipment, and by acceptability how much the subjects liked using the equipment for the task.

These two main dimensions are still used for both the description and the evaluation of usability. Definitions along these lines have been applied in the European research project COST 212 (Human Factors in Information Services) as well as within the European standards institute of ETSI. Internationally, similar although not identical definitions are found in ISO 9241-11 "Guidance on usability" (1996).

3.1 Usability concepts

With the above background, measures of usability may be assumed to be of two kinds:

a) *Performance measures*, which are "objective" measures or observations of user behaviour and are focused on task performance, i.e. how well the users can achieve a specific task.

b) *Attitude measures*, which are "subjective" measures of the users' opinions of working with the system, i.e. how much they like using the system.

Performance measures include both *Effectiveness* (e.g. error rate) and *Efficiency* (e.g. performance time), which are considered to be independent factors, in the sense that they can vary without necessarily being correlated. The following definitions are given in ISO 9241-11:

Effectiveness: The accuracy and completeness with which users achieve specified goals.

Efficiency: The resources expended in relation to the accuracy and completeness with which users achieve goals.

Attitude measures include concepts such as satisfaction and acceptability. The attitude measure is rather more complex than performance – it should be seen as a conglomerate of underlying dimensions, which are all related to the user's perception and experiences of the system when it is used for solving a certain type of task. Typical examples of such dimensions are *Satisfaction, Acceptability, Social presence, Comfort, Enjoyment, Control,* and *Aesthetic appeal.*

It should be observed that when usability is measured, the outcome is always specific for the kind of *user* involved, the *task* to be solved, and the context and *environment* in which the test has been carried out. ISO has furthermore introduced the concept of *Context of use*, defined as:

Context of use: The users, goals, tasks, equipment (hardware, software and materials), and the physical and social environment in which a product is used.

Consequently, the concept of Usability cannot be generalised unless it has been verified how much it is affected by variations in the parameters user, task and environment. Therefore, the additional concept of *Flexibility* has been introduced as a measure of these effects, in the sense that the less usability changes over variations of any of the three parameters, the larger is the flexibility of the system in question.

Finally, the change in Usability as measured over a number of trials with test subjects representing users, is a measure of *Learnability.* Requirements for learnability need not be the same for different user groups: for infrequent users, e.g. users of public telephone terminals, it is often important that the use is successful even at the first trial, while for office equipment it might be better to aim at a high efficiency for professional users, which may be achieved only after a relatively long training period.

3.2 Usability and telecommunications

The ISO 9241 definitions of usability components are of a general nature, but the standard in fact addresses particularly the ergonomical requirements for office work with visual display terminals. For telecommunication terminals similar performance criteria may be applied; however, the evaluation of user attitudes may need further consideration when usability concepts are applied to telecommunication systems and services.

The evaluation of user attitudes in a usability test is considerably facilitated by identifying in advance those dimensions of the acceptability/satisfaction concept which have the best differential power, i.e. which most clearly show the difference between systems under test. Opinion rating scales which give the same score for systems that are perceived as significantly different are not effective as test tools. In the same way, rating scales which all give the same score for a specific test object are probably redundant and therefore less efficient. In this respect the methodological situation is the same as for the evaluation of speech quality.

For person-to-person telecommunication, the Communications Studies Group have shown that the concept of acceptability comprises an important dimension which may be called *Social Presence*. This factor is of major importance for the comparison of different communication systems and is therefore a suitable basis for the design of questionnaires. Systems having a high degree of Social Presence are judged as being "warm, personal, sensitive and sociable".

Another factor of particular importance for the evaluation of terminal equipment has been called *Aesthetic Appeal*. This dimension is positively correlated with adjectives such as colourful, large, spacious, beautiful and interesting. However, Aesthetic Appeal is also strongly related to the Social Presence dimension and may even be confused with it.

"Aesthetic appeal" does not necessarily refer to any aesthetic or artistic value but rather to the appearance, the perception of shape and colour and even the tactile impression of the equipment representing the human/machine interface. No doubt these are components which have a significant influence on the users' attitudes towards a telecommunication system, particularly its "acceptability". There are in fact numerous examples of products (be they telephone sets or hearing protectors for work in noisy locations) where the user has preferred the type having the more elegant light-weight design over other types with better performance but with a less appealing look and heavier weight. These examples bring out the fact that it is the users who judge which product is most usable by their personal weightings of performance against satisfaction or acceptability, not any "objective" external judges with their expert opinion.

Further dimensions which have proven useful for explaining the results from attitude questionnaires especially for the use of conference telephone systems are, according to CSG (1974):

- *Convenience*
- *Privacy*
- *Control of conversation*
- *Polarisation between participants*
- *Concentration*
- *Time pressure.*

Another set of dimensions for the assessment of satisfaction has been proposed by the Heinrich Hertz Institute in Berlin in a contribution to the ETSI work on usability evaluation. They found in tests on some telecommunications services, particularly multimedia information systems, that the following three factors can be used for describing satisfaction in a communication situation:

- *Effort* versus *Ease of use*
- *Agitation* versus *Relaxation*
- *Boredom* versus *Enjoyment.*

The Effort/Ease-of-use dimension is positively correlated with such adjectives as easy, simple and restful, Agitation/Relaxation with relaxed and calm, and Boredom/Enjoyment with entertaining, animating and enjoyable.

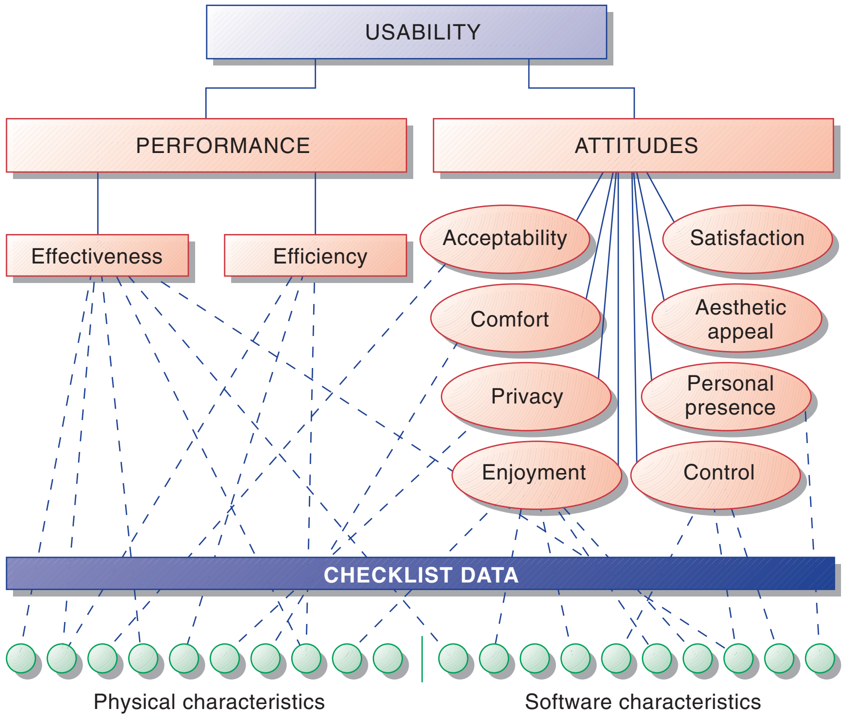


Figure 12 Usability and its components. The model has three layers, where the bottom layer consists of the physical factors that have an impact on the performance and attitude components, which are mutually independent but may depend on the same underlying factors. "Checklist data" are recommended values for the bottom layer characteristics in order to achieve an acceptable level of usability

3.3 A conceptual model

Similarly to the model of speech quality presented in Figure 2, the components of usability are expected to depend on a number of underlying factors, as illustrated in Figure 12. The two dimensions of usability: *performance* and *attitude,* should be considered as orthogonal, i.e. as mutually independent factors. Instead of being interdependent they may both depend on the same underlying factors but to a different degree. They may also depend on totally different factors. These factors could be any characteristics of a product or service that might have an influence on the usability components, including "soft" factors such as the dialogue structure in a voice response service.

The assumption of orthogonality means that a product or service can get a high score on one component and a low score on the other component as the result of a systematic evaluation. An example: one type of conference telephone which in a usability test was proven to be effective and efficient for solving tasks by telecommunications that otherwise would require a personal meeting, at the same time had a low acceptability for the users, because it was perceived as bulky and uncomfortable to use. In the same test, another type of telephone was preferred because it was considered to be easy to use, without actually being a very efficient tool for the intended task.

If it were possible to estimate the weighting coefficient that an individual user would attach to the usability components, a total usability measure could be obtained that in the example above might turn out to be about the same for both telephones, in spite of the obvious difference in performance and user satisfaction. For the same reason as for speech quality, splitting up usability into relevant components facilitates the evaluation of a product and gives detailed information about the strong and weak points of the product.

Comparing the conceptual model for speech quality with the usability model, it appears that intelligibility could be said to correspond to the performance dimension, while naturalness is related to attitude measures comprising satisfaction, acceptability, etc. Intelligibility can be assessed by means of "objective" tests, where the performance of the users is observed from the outside. The proportion of correctly understood words is a measure of goal achievement and may be considered as the "effectiveness" as defined by ISO. The number of words understood per minute is a measure of message rate and may be said to be the "efficiency". However, message rate is hardly a relevant measure for characterising a telephone system, because this rate is much more dependent on the users' behaviour in a task solving situation than on the transmission characteristics of the system.

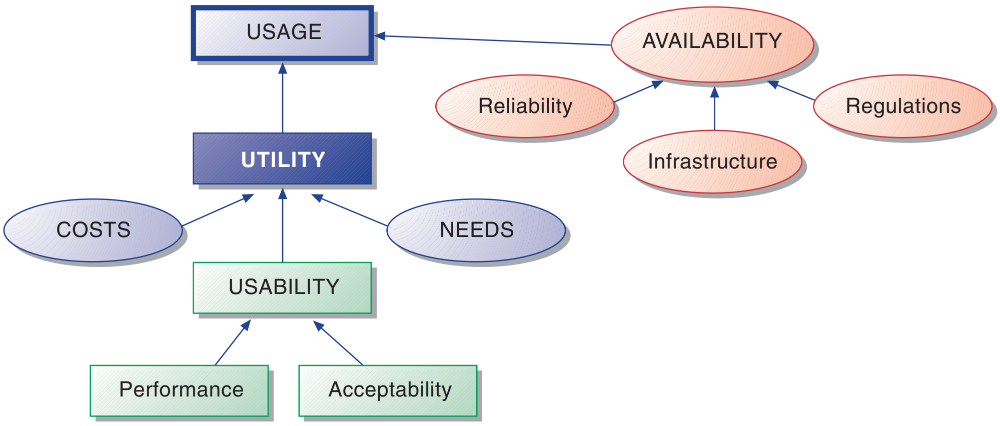


Figure 13 A model of Usage, Utility and Usability

Naturalness, on the other hand, can only be assessed by subjective tests, where the users have to give their opinion on specific rating scales. It is therefore comparable to the "satisfaction" component of usability. The difference is merely that while naturalness may be considered to be one-dimensional, the attitude components of usability consist of a rather large number of independent dimensions which all have to be rated separately, if a complete evaluation of the usability of a product is aimed at.

Therefore, the principal difference between the concept of speech quality and the even more complex concept of usability is that it is hardly meaningful to try to estimate usability directly on a single scale in a test, while it is often possible to obtain a relevant overall speech quality measure without going into further detail.

3.4 Utility and usage

It should always be kept in mind that usability is a quality measure that does not include the costs for providing the product or service. It relates only to the *ergonomic* quality of using the product. Even if "efficiency" involves the resources expended, which includes the physical and mental effort as well as the time spent by the user for task completion, it should not include the financial costs for providing or using the product. Instead, the concept of *Utility* should be applied, as illustrated by Figure 13.

Utility may be seen as the relative value a user attaches to a product by taking the costs into account together with the benefit of using the product for a specific purpose.

One advantage of such a conceptual model is that it clearly demonstrates the role of usability *in relation to* the cost and the user's need. It is a necessary but not sufficient condition that usability reaches a certain minimum level. However, shortcomings in the level of usability may be compensated by either a strong need for using the product or by a low cost, or both. Maximum utility is obtained by the highest possible usability together with a sufficiently low price in relation to the needs of the user. With this model, utility comes down to zero even for highly usable products, if the consumer does not need them, or if the price is out of reach.

With a given utility of a product for a relevant user group, there will be a certain demand which finally leads to some degree of usage of the product. However, there are further conditions which have an influence on the usage. One group of such additional factors may be called Availability. Especially within the area of telecommunications, it is a necessary condition to have access to a reliable network and to be able to reach a sufficient number of other users of the same service. If all conditions are fulfilled when a new service is introduced and the usability and utility is high enough, a rapidly growing usage is possible and may make the service an immediate success. Eventually, if a comprehensive pool of empirical data is available (e.g. checklist data as shown in Figure 12) and the relations between the factors in a structural model are reliably known, it may become possible to predict the success of a new service before it is introduced.

4 Service quality

Originally, "Quality of Service" dealt exclusively with transmission performance and reliability, particularly in ITU-T documentation. However, recently more and more attention has been devoted to *perceived* performance, which means that customer satisfaction is coming into focus. In ITU-T Recommendation F.70 [10] the main criteria of service performance from the users' viewpoint are:

1) ease in establishing a connection
2) retention of the established connection
3) satisfactory transmission quality
4) integrity of billing.

Applying a usability perspective to service quality may be facilitated by adopting the classification of tasks which has been developed within the European RACE/GUIDANCE project. A distinction is made there between *goal tasks*, which are the tasks that the user wants to perform, and the *enabling tasks*, which are what the user must do to create a state which enables the goal task to be performed. The first point in F.70 deals with the enabling task, the second point with reliability, and the third point with the goal task. The fourth point is a different issue, which today might be generalised to deal with service integrity and security.

For a telephone service, usability aspects can be applied both to the access phase (enabling) and to the speech communication phase (goal) which generally make up the service. For both phases, usability targets can be set up and evaluated. The ease in establishing the connection is very much dependent on the procedure for setting up the call, which may involve access codes, verbal dialogues with an operator or a speech recogniser, payment by coins or cards, etc. User instructions also play an important role for the ease of access, as well as the design of the physical interface in the terminal.

Therefore, the assessment of service quality should always involve a usability evaluation of the complete access phase of the service by measuring user satisfaction as well as performance (e.g. percentage of achievement at first trial, or error rate, on the one hand and time for completion on the other hand).

The communication phase should be dealt with separately. For evaluation of the quality of communication by speech, the division of speech quality into perceptual factors, which may be studied independently, is a useful approach. Presently, there is a clear tendency to apply the task oriented methodology implied by the usability model also for the communication phase.

As an example, according to ITU-T Recommendation P.85, the quality of a speech synthesis system should be assessed by application of the system in a simulated service where the subjects perform specific tasks in a given scenario, e.g. a railway traffic information service. In the test, the performance of the subjects is recorded, and the degree of acceptance of the voice service is assessed by asking a question like *"Do you think that this voice could be used for such an information service by telephone?"* After that, more specific speech quality rating scales are applied, such as: *Overall impression, Listening effort, Articulation, Pronunciation, Pleasantness*. The main point is that the synthesis system is not rated per se but in a service application. In this way speech quality can be made an integral part of service quality.

A further factor that may affect perceived service quality – apart from all the factors contained in the usability concept – is the users' expectations. A user's opinion of a service is no doubt coloured by expectations about what can or cannot be achieved by it and how the possible benefits compare with other services. For example, it appears that users of mobile telephones have a more lenient attitude to the speech quality of mobile services in comparison with the fixed telephone network services, because the advantage of mobility makes the user judge a degradation in sound quality less seriously. This factor has been taken into account in a model for the calculation of speech quality recently developed by ETSI [7].

5 Conclusions

It has been shown here that the concepts of speech quality and usability both stand on a similar basis and can be described by parallel models. Speech quality has two basic components: one is related to speech as a means of communication and describes the performance of a telephone system when transmitting semantic information, and the other is related to conveying personal and emotional characteristics imbedded in the voice of the speaker. Similarly, the usability of a service has an "objective" dimension, describing goal achievement and efficiency, and a "subjective" dimension, describing user satisfaction and acceptability.

Both models can be integrated in the evaluation of service quality. They are applicable both with respect to different evaluation mehtods and for the assessment of service quality by calculation methods. The models may also accommodate additional factors such as reliability and availability as well as user expectations in the context of use and, ultimately, also the cost of using the service in relation to the benefit for the user.

Descriptive models based on carefully defined fundamental concepts are a good help for understanding why some telecommunications services succeed and others fail. They may furthermore be considered as useful tools for research and development which can promote the design of both effective and more user oriented services.

References

1 Gleiss, N. Desirable sending frequency response of telephone sets. *Tele Engl. Ed.,* 1, 1989.

2 Gleiss, N. Sound transmission quality : the influence of signal-to-noise ratio and bandwidth. *Tele Engl. Ed.,* 2, 1972.

3 Gleiss, N. Preferred listening levels in telephony. *Tele Engl. Ed.,* 2, 1974.

4 Gleiss, N, Goldstein, M. Loudness of speech for monaural and binaural listening. *7th FASE Symposium,* Edinburgh, 1988.

5 McGee. Semantic components of processed speech. *Journal of Acoust. Soc. Am.,* 1964.

6 ISO. *Ergonomic requirements for office work with visual display terminals, Part 11 : Guidance on usability.* ISO 9241, 1996.

7 ETSI. *Speech communication quality from mouth to ear for 3.1 kHz handset telephony across networks.* ETSI/ETR 250, 1996.

8 ITU-T. *A method for subjective performance assessment of the quality of speech from voice output devices.* ITU-T Recommendation P.85. Geneva, 1994.

9 ITU-T. *Transmission impairments.* ITU-T Recommendation G.113. Geneva, 1996.

10 ITU-T (CCITT). *Evaluating the quality of the international telex service.* ITU-T Recommendation F.70. Geneva, 1988.

11 ITU-T. *Models for predicting transmission quality from objective measurements.* Supplement 3 to ITU-T Series P Recommendations. Geneva, 1994.

Norman Gleiss is the former Human Factors Manager (now retired) at Telia Research, Sweden. He is also the former Chairman of ETSI Technical Committee Human Factors.

e-mail:
norman.a.gleiss@stockholm.mail.telia.com

Some current issues of Human Factors in telecommunications

KNUT NORDBY

Scope of the review

Human factors in telecommunications covers such a large and diverse field that it is quite impossible to cover in any detail all its varied aspects in a review like this. However, I will here try to give an overview of some current topics for those who are new to the field, but readers who wish a more thorough treatment must look to the specialist literature. The most redeeming feature of this overview is probably that it is reasonably up to date since much of it is based on contributions to the *16th International Symposium on Human Factors in Telecommunications (HFT'97)* held in Oslo, Norway, 12–16 May, 1997 [1].

This review is meant to present an account of what is currently going on within the many fields of human factors in telecommunications research in various laboratories around the world. To simplify the overview, the different topics have been grouped together under eight broad headings, corresponding roughly to the main sessions of the *HFT'97* symposium. These are:

1. User interfaces – simple and complex
2. Speech interfaces
3. Video communication and multimedia
4. Design for all – support for elderly and disabled
5. Group working support services
6. Product development
7. New services, user requirements and marketing
8. Convergence – integration of telephony, computing, Internet.

In a review like this it is difficult to avoid technical terms. However, I will try to keep the use of very special technical terms and abbreviations to a minimum and to explain them as they appear. A full list of references is provided at the end of the review, but those who want to go into any depth are advised to consult the specialist technical literature.

1 User interfaces – simple and complex

The telephone user interface has remained nearly unchanged for a very long time, basic call set-up procedure being virtually the same today as 80–90 years ago when Strowger automatic telephone exchanges came into use. This means that three to four generations have grown up with and learned this procedure. As a result of the emergence of new technologies, with the merging of telecommunication and information technologies, and the competition made possible by deregulation and privatisation of operating companies, a multitude of new services have grown up requiring new interfaces. Most notable are the interactive voice telephone services where you are asked to press number keys to input your choices, or with voice recognition, where you may speak your choices. Since most users have grown accustomed to interact with *people* via the telephone, interacting with the impersonal *telephone system* itself is a new experience to most people and has to be learnt by the common user. Unfortunately, the design of these new interfaces has been aimed more at well educated, computer literate, young men, than at common users, including people of all ages, all education levels and all levels of mastering.

The oldest and most important form for feedback to users of the telephone system is the *line signals* or 'tones' you hear in the handset. Most users have now grown accustomed to their 'local tones', but recently a number of new tones have appeared, confusing the users. If you travel abroad or phone another country, you will often encounter new and strange tones, whose meanings are not always obvious, leading to more confusion and uncertainty. There is obviously a strong need for standardisation and harmonisation, but should we keep the 'hard core' of traditional tones or should we look at the information tones from a different angle and design new tones which convey the meaning and degree of urgency in a more universal way. Traditions are hard to change and many people are loath to give up their well-known tones to some strange sounds imposed upon them by a distant and impersonal international agency.

It is remarkable that so many new features are currently implemented in new telephone instruments without the necessary attention to human factors of the user interface. New functions abound with buttons marked with non-standard, unintelligible symbols, new user procedures breaking with established practices, bad ergonomics with no tactile feedback from keys, illegible symbols and displays, and badly designed handsets which are strenuous to hold for any length of time, etc. The manufacturers maintain that it is not worth putting any resources or efforts into the ergonomics of telephones, which are looked upon as cheap consumables rather than investments. It is hard to see the logic of implementing many features which most users will never use, except as a marketing ploy. Badly designed telephones, however cheap they may be, will always be a liability to the users, and may ultimately lead to missed business opportunities, which no company can afford. When new telephones were introduced at the University of Oslo, hardly any users could perform even the simplest actions like call transfer, but had to ask the caller to call back to the operator and ask for the other party. This was not merely a learning problem, but the result of indifferent interface design. Most staff just would not take the effort to look up the procedure in the instruction manual every time they had to transfer a call, and the operation was too complex, involving arbitrary codes, to remember if you did not perform the operation regularly. Also, there was no feedback from the system to inform you if you had performed the operation correctly, often leaving the unfortunate callers on hold indefinitely.

1.1 Audible tones in public telephone networks

The most important feedback to the users of the public telephone system is the various tones you hear in the handset (i.e. the *line signals*) such as the *dialling tone*, the *ringing tone*, the *engaged tone* and the multitude of other more or less intelligible tones introduced by various operators.

A comprehensive study of the human factors aspects of information tones in public telephone networks was carried out by D. Anderson, A. Ferris and W. Mellors, which is published as ETSI Report ER 145 Part 1 and 2 [2,3] (ETSI = European Telecommunication Standardisation Institute). The work was mandated by the CEC (Commission of the European Communities) to investigate the current and future use of information tones in the public telephone network with the view of harmonisation in Europe to avoid the possibility of cross border misinterpretation and confusion of tones by users from different countries. There is an urgent need for harmonisation and simplification since an initial survey by Gagliardi had found that at least 228 different information tones were in concurrent use in Europe alone! (see [2,3,4,5,6]; and companion article by Anderson & Ferris, *Feedback tones in public telephone networks: Human Factors work at ETSI*, in this issue of *Telektronikk*).

Anderson, Ferris & Mellors have developed a three-dimensional user-goal model (Figure 1) where each axis represents a semantic pair:

1. Progress – Failure to Progress
2. Prompt for Action – No Action Required
3. Interaction with System – Talk Phase.

Based on this model and functional definitions of the most basic line signals, Anderson *et al.* devised some new tones to match the various semantic characteristics defined in the model: such as the *degree of urgency, the phase of the call*, and *call progress*. It is thus possible to propose basic information tones with high human factors relevance to replace today's tones which were based on technical possibilities in the early telephone networks rather than on human factors.

Each new tone, as well as some existing commonly used tones, have been subjected to an 'audit' based on the functional definitions, best fit with the user-goal model, and on technical characteristics. As a result if this audit, Anderson *et al.* propose six new tones for Pan-European standardisation: *dial* tone, *ring* tone, *busy* tone, *congestion* tone, *special information* tone, and pay tone [2,3,6]. The tones have not been tried on common users yet, but knowing the users' general resistance to any change, their reactions to the new tones should be interesting.

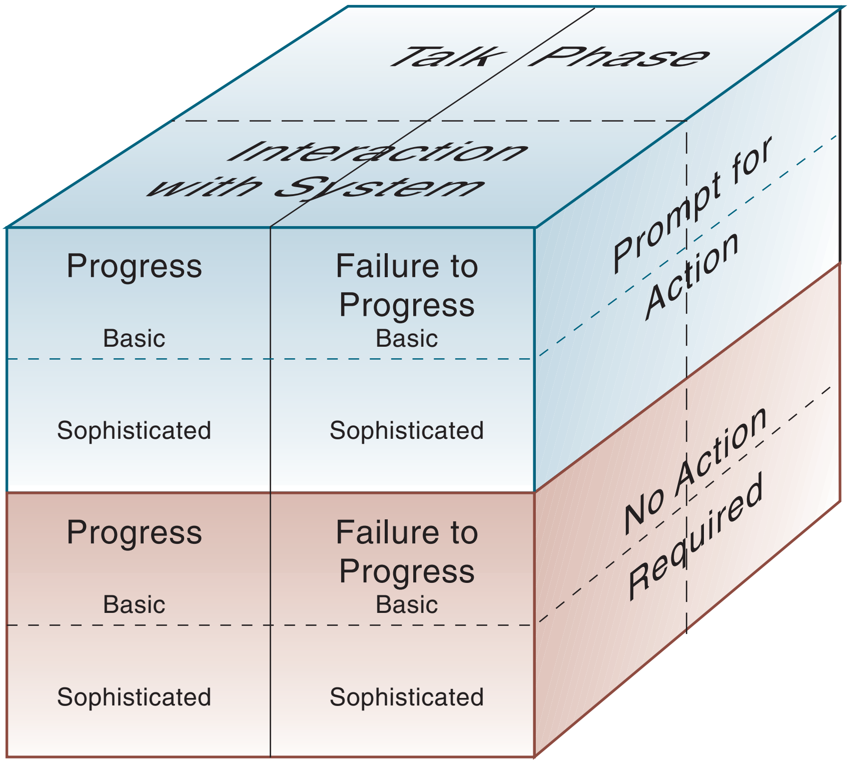


Figure 1 Three-dimensional user-goal model for categorising network information tones in public telephone networks along three dichotomy dimensions. These are: Progress ↔ Failure to progress; No action required ↔ Prompt for action; *and* System interaction phase ↔ Talk phase. *Both existing and proposed information tones can be characterised along these three dimensions (after [6])*

1.2 Is it worth having ergonomic telephone user interfaces?

A recent problem for many users is the many functions found in modern telephones, which no-one seems to use. Why are user interfaces so poor; why have the human factors specialists had so little impact on product development; and it is really worth the cost and effort to improve the user interface of telephone functions no-one wants to use.

In a critical paper, Helmreich [7] from Siemens, Germany, focuses on these questions. Initially, Helmreich observes that to sell a product one must advertise its many functions. The customers must believe that the more functions a product possesses, the greater benefits will be derived from buying it. Common experience, however, has shown that most novel features are rarely, if ever, used – over-featuring seems to go hand-in-hand with under-use.

Sadly, Helmreich points out, the ergonomic aspects of many telephones in widespread use do not even meet minimum user requirements. Low contrast displays are difficult to read, very minute buttons and dialling keys are often 'spongy' (i.e. have no clear tactile feedback), operating procedures are inconsistent and markings on operating controls are either incomprehensible, illegible, or both.

Helmreich further maintains that the most deciding factor for telephone manufacturers are costs: "Ergonomics is all right, providing it doesn't cost any extra". He then goes on to show that the use of ergonomics in telephone design does, in fact, pay off in real money. He compares a set of newly designed telephones from Siemens, where the best ergonomic principles have been used, to some conventional telephones, and is able to show in a series of analyses that the user-friendly telephones do, indeed, give an economic benefit over the conventional telephones.

Helmreich first looks at the *learning phase*, where he can show a saving during the introductory learning phase of between DEM 12.50 and DEM 31.25 per person. Everyday use, however, shows even greater savings. Helmreich shows that in an office with a PABX with one thousand lines the savings from the use of the new telephones (especially in increased use of call transfer and conference calls) add up to 75 seconds per week, giving an overall saving of up to DEM 206 per week for each telephone.

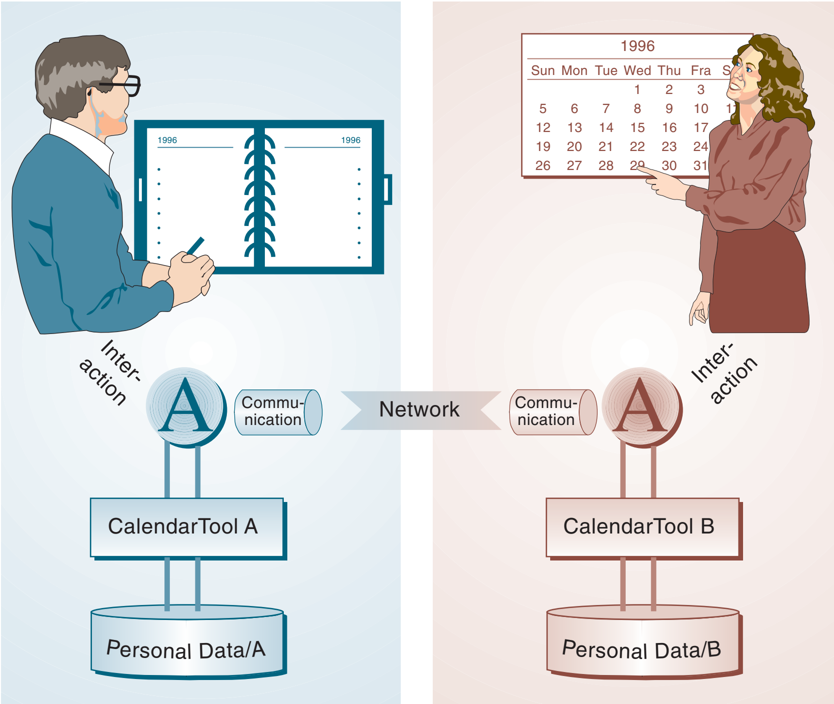


Figure 2 The architecture of the appointment scheduling system: The users work with calendars they are used to. Intelligent agents, acting as wrappers, communicate with each other via standardised protocols, which allows integration of multiple calendar systems (after [8])

Helmreich also looked at factors that cannot be quantified. By having telephones that are easier to use, the running of the daily business will be more efficient. The telephone has become such an important instrument in the daily running of businesses that mistakes or lack of confidence in e.g. transferring a call, can, in the worst case, lead to loss of customers and revenue. Such factors cannot easily be quantified, but may have a considerable effect on operating results and productivity.

The savings found may not look impressive, but an average of DEM 200 a year per employee, however, will quickly add up to savings running into many tens of thousands of marks a year in larger companies. Paying, say as much as DEM 100, extra for a more ergonomic telephone will, thus, pay for itself in only six months. In conclusion, Helmreich maintains that if it is possible to demonstrate economic benefits for manufacturer and customer, there is a greater chance that ergonomic design will be used when developing a new product. If this is done as an integral part of the development process, he believes that there will hardly be any increase in development costs [7].

1.3 Intelligent agents as appointment managers

Time management has always been a problem, both in work and leisure. A variety of support tools for time planning and management have been devised in recent years, first paper based, but increasingly in electronic form. The task of scheduling an appointment is both time consuming and requires a high degree of co-ordination; especially when more than two participants are involved; each of which may have a busy schedule. Traditionally, several iterations, face-to-face, via telephone, fax or e-mail were necessary to find a time acceptable to all parties. The likelihood that not all participants are on hand at the same time complicates further the joint scheduling process.

With the increasing use of electronic time management tools, the process of scheduling appointments can be considerably simplified or even fully automated. Dieterich & Steiner [8] of Siemens, Germany, show that autonomous 'intelligent agents', acting on behalf of their users, can take over the tedious task of negotiating appointments. Intelligent service nodes in telecommunication networks will not only enable agents to communicate with each other, but also to dispatch messages via fax or e-mail, or convert them into interactive voice messages sent to telephones or Personal Information Communicators.

The basic idea put forward by Dieterich & Steiner is to utilise existing infrastructure (telephone and GSM networks) for easy accessibility, widely available electronic calendar tools for wide availability, and a soon to be standardised open communication and negotiation protocol [9] (see Figure 2).

Dieterich & Steiner used the Multi-Agent Interaction and Implementation Language (MAI²L®) developed by Siemens for implementing intelligent agents in C® on Unix® and Windows32® platforms as part of the Multi-Agent Environment for Constructing Co-operative Applications (MECCA®). The scheduling capabilities of such agents are still limited to finding a time slot that is free for all. It will not consider constraints such as the time it takes to travel to the meeting venue or various personal priorities. However, Dieterich & Steiner think that a wide variety of techniques developed for distributed appointment scheduling, such as automated learning of user preferences, priority management and conflict resolution, can easily be incorporated into this framework. Standard human-computer interaction techniques, such as natural language understanding, should also be of use for human-agent interaction.

2 Speech interfaces

The most direct human information interface has always been speech, which is a genuinely human ability acquired by virtually all the members of the community. Until recently, however, the technology necessary for consistent recognition of the human voice has not been available, and alternative means for interaction with telephone systems have been developed. With the introduction of DTMF (= Dual Tone Modulation Frequencies), 'touch-tone' to some people, it became possible to utilise the telephone for other things than conversations or one-way information services (e.g. the speaking clock, news, weather reports, etc.). By being able to give input to the system, users could engage in interactive services without the intervention of a human operator. One of the first interactive services to be offered to the general public was a service allowing people to check their own bank accounts. This is now 'tele-banking', allowing people to transfer money from one account to another or pay bills via their home telephone without the need of any extra equipment. Another early service was provided by book clubs, allowing their members to cancel or order books via the dialling keypad of their home telephone.

Today, such services have become so common that it is hardly possible to phone any agency or large business concern without being greeted by: "If you want this, press 1, if you want that, press 2". For most people the procedures for using many of these systems are too complicated since you have to remember service codes, account numbers, item codes and key-mappings, leading to 'slam-down' and under-use.

The handling of error recovery in most services leaves much to be desired – most people who make an error just hang up. Feedback is another neglected issue. Even minute differences in the wording of prompts and feedback messages can mean the difference between success and ruin of a service. In an interactive service introduced in UK, the users were asked to "Press the *harsh* key", which lead to users not understanding how to proceed. Although the term *hash* is the 'correct' name for the # symbol, it is generally known as the *square* in UK, and when the message was changed, the service was a success. Inconsistent naming of the two symbols leads to confusion. The * symbol is usually called *asterisk* or *star*, or their equivalents in different languages, but the # symbol has many different names, even within one language, e.g. *harsh*, *square* (UK), *pound* (US), *number sign*, *sharp* (music), *double cross*, *little gate*, *fence*, to mention only a few. More important, though, is the inconsistent use of the two symbols. The * key is usually used to 'initiate' a service, as 'separator' between different inputs or as 'affirmative' key, but also, and inconsistently, as 'delete' (backspace) key, while the # key has been used both to 'initiate' services, to 'confirm' choices and as 'delimiter' key to terminate a service. Even the standards published by ETSI (European Telecommunication Standards Institute) and the recommendations from ITU-T (International Telecommunication Union) show inconsistencies, and harmonised standardisation is urgently needed.

The alternative to using the dialling keypad is to use voice recognition allowing the user to speak the various commands and inputs to the system. The problems associated with making a good and simple user interface with this technology is even more daunting and much effort has gone into this during the last years, as demonstrated below.

The inherent impreciseness of colloquial speech makes it difficult to design a dialogue that will work smoothly for different users in different contexts: It reminds me of the Norwegian folk tale about a deaf man who, seeing the sheriff approaching, carefully thinks out clever answers to the questions he expects, and the hilarious dialogue which follows when the questions do not match his replies. Speaking to a telephone system still puts many people off, especially if the dialogue is jerky and unnatural. Speech recognition as input to interactive telecommunication services is surely here to stay, but we have to design its dialogue properly for it to survive in the marketplace and make revenue.

The following sections illustrate various attempts to design interactive telecommunication services and to compare different dialogue designs, both in the laboratory and in real-life settings.

2.1 A common keypad interface for voice-response services

Different telephone voice-response services utilising the dialling keypad (e.g. messaging, interaction and information retrieval) employ different functional interfaces, leading to users having difficulties when switching between services. Goldstein [10] of Telia Research, Sweden describes a common functional keypad interface proposed by Telia Research, and how it was evaluated in three voice-response services.

By using the 'Tape Recorder' metaphor together with the auditory *AlphaWheel* paradigm and the 'Constrained Output Set', a universal navigation concept, called a common functional keypad interface, was developed by Telia for use with a voice-messaging service (*MobilSvar*). The basic question Goldstein asked was

*Figure 3 A common functional keypad interface proposed by Telia for all kinds of interactive voice response services: Keys 1–5 for menu options; key 6 = change/erase; keys 7–9 for the 'tape-recorder' metaphor (key 7 = fast rewind, key 8 = stop/start, key 9 = fast forward); and keys * 0 # for decision functions (* = undo/cancel, 0 = help, and # = finish/execute) (after [10])*

if this interface could also be used in an information retrieval service.

The Telia functional keypad interface conceptually divides the standard telephone dialling keypad into three areas (see Figure 3): Menu options (1 2 3 4 5), Change/erase (6), the 'Tape Recorder' metaphor (7 8 9) and Decision Functions (* 0 #).

The 'Tape Recorder' metaphor involves mapping the keys of the standard music cassette player onto the 7 8 and 9 keys so that the 8 key plays the message, the 7 key goes back (rewind) and the 9 key goes forward (fast forward). Functions are toggled on and off by repeated key presses.

The auditory *AlphaWheel* paradigm is a software application for entering alphabetic input in any language from a standard telephone numeric keypad. The system recites the letters of the alphabet "A, B, C, ..." etc. and when the first letter of the desired name has been reached the user selects (or 'clicks') it by pressing the # ('Execute') key. The system (if necessary) recites the alphabet again and the second (and third) letter of the desired name is selected. By using the 'Tape Recorder' metaphor keys, the user can move quickly through the alphabet (9 'Forward' +5 letters and 7 'Rewind' –5 letters at a time). When enough letters (usually two or three) have been selected, the 'Constrained Output Set' will recite a short (3-5) list of names beginning with the selected letters. The subject can then choose the desired name by 'clicking' it with the # key (if there is only one name in the list beginning with the first selected letter it will be selected immediately).

In the user trials a Telia bank-by-phone service, using the common functional keypad, was significantly more efficient than a comparable service (*Datasvar*) on the Swedish market. Two public information retrieval services (a Ferry Service with 210 port names, and the Greater Stockholm Public Transport Service with 350 station names), both which use a digit code paradigm, where implemented with Telia's common functional keypad interface. Comparisons between the original services and Telia's implementations gave diverging results; the Telia ferry service was significantly more efficient than the code-dependent service, while the original code-dependent Stockholm Public Transport Service was significantly more efficient than Telia's implementation. However, if flexibility and learnability are also taken into account, the results are in favour of Telia's functional interface since no code manuals are required. By offering customers one common functional interface for all interactive voice services, transfer from one service to another is made easier. The functional keypad interface is cheap to implement both from a user's and operator's point of view, as it is based on available technology and is terminal and language independent. It is also supported by mobile telephones.

The proposed common functional keypad interface does not conform to current ITU-T or ETSI standards for definitions of telephone keypad key functions. More human factors assessment and extensive full-scale user trials will be needed before standardisation of the functional keyboard interface can be undertaken.

2.2 Speaker verification security in automated services

Procedures for security checks in automated telephone services, such as home banking or home shopping, require particularly careful design of the dialogue between system and user. This is because of the 'binary' nature of the access decision; you are either accepted or rejected. There are two options for the designer, either to use *user-specific knowledge*, such as passwords and PIN-codes, or to use biometric technologies, which rely on *specific physical attributes of an authorised user*, such as speaker recognition, fingerprints or iris identification. Of these, automatic speaker verification holds distinct potential since security would not only depend on knowledge of passwords or digit codes, but also on user-specific voice attributes. Unlike ordinary dialogue design, where errors can be dealt with by well-designed error recovery strategies, verification of user authorisation results in either access or rejection. In the case of rejection, system designers can either provide immediate human-operator backup, or introduce repeated automatic verification cycles, with the risk of increased 'insult rate' should the user fail to gain access after repeated verification sequences, or terminate the transaction.

Schmidt*, Ironside*, Jack* & Stentiford# [11], *Centre of Communication Interface Research, University of Edinburgh, Scotland, and #BT Laboratories, Ipswich, England, performed two experiments; one on security data types and dialogue messages, and one to determine 'insult rate' and verification completion. The experiments addressed the question of how to proceed in dialogues with front-end speaker verification to design 'best practice' hand-over strategies for transferring a user to a human operator should difficulties arise.

The results showed a general (and expected) decline in user attitude with increasing 'insult rate'. Although there was no significant difference between successful completion by machine or human operator on the second attempt, analysis of individual statements point to a preference for the second attempt being carried out by machine rather than suggesting an immediate hand-over to a human operator. These results may be explained by a different voice being introduced by the human operator. This requires heightened concentration, increasing the cognitive load.

First and foremost, their findings show that users have strong opinions towards these issues and, consequently, that the speaker verification procedures in automated telephone services need to be carefully researched and designed.

2.3 Experience with 'Voice Server' for overseas calls

A voice recognition 'Long Distance Voice Phone Calls Server' for managing overseas telephone calls for callers who had the necessary information and had no need for other services had been introduced in Colombia in 1995 as a complement to telephone operators. Charry, Zuñiga & Mejia [12] of ITEC-Telecom R&D, Bogotá, Colombia, studied the user reactions and user acceptance of this system.

Due to the diversity of Spanish dialects spoken in Colombia, the varying levels of technical training, and the need to speak telephone numbers one digit at a time, the system was not very well received. Initially in the Bogotá region, 45 % of the callers hung up when encountering the new service, 20 % switched to a human operator, 16.5 % failed to complete the call and gave up, and only 18 % completed successfully the attempted calls using the Voice Server (0.5 % of the calls could not be completed for various technical reasons).

Even after having been in use for some time and after changes and improvements to the system, these numbers did not change dramatically: (38 %, 38 %, 8 % and 15.5 %, respectively). Actually, the number of successfully completed calls *fell* from 18 % to 15.5 %, but more people switched to a human operator. Charry *et al.* propose various cultural, linguistic and ergonomic explanations to this:

- The system was introduced abruptly and there was a lack of publicity.
- Colombians do not feel any need for using such technology and are even afraid of using it, although they think that the company does not provide sufficient operators to provide proper service.
- The majority of users had never used any type of voice recognition technology before and therefore rejected the service.
- The demand to speak long telephone numbers one digit at a time was quite unnatural to most people, and callers had to write numbers down to be able to read them out in this manner.
- Colombians have strong feelings of regional belonging, and not being greeted in their local dialect put many off using the system.
- Spanish synthetic speech is not yet well developed, and many users found the prosody strange and rejected the system.
- Spanish voice recognition is in an early stage and has a very limited vocabulary.
- The interface did not provide sufficient help and information to instruct callers in using the system correctly.
- Any additional help, however, is too time consuming, frustrating the callers' patience.
- The system is not 'dynamic' enough; many speakers interact like in a conversation, starting to speak before they hear the prompt tone.
- It is not possible to recover from an error, and you have to hang up and try again.

For the Voice Server to be viable, Charry *et al.* feel that it is important that the dialogue is natural, free flowing and that regional dialects are respected. Some of the problems pointed out will be solved in the near future. Also, voice synthesis and voice recognition is improving in most major and even some small languages. In cultures where high tech is recognised, resistance to new services will abate. But this does not give operators the freedom to introduce any bad service. The demand for high quality interactive services will rise, and as competition will give a choice between similar services, users will choose those which are simple to use.

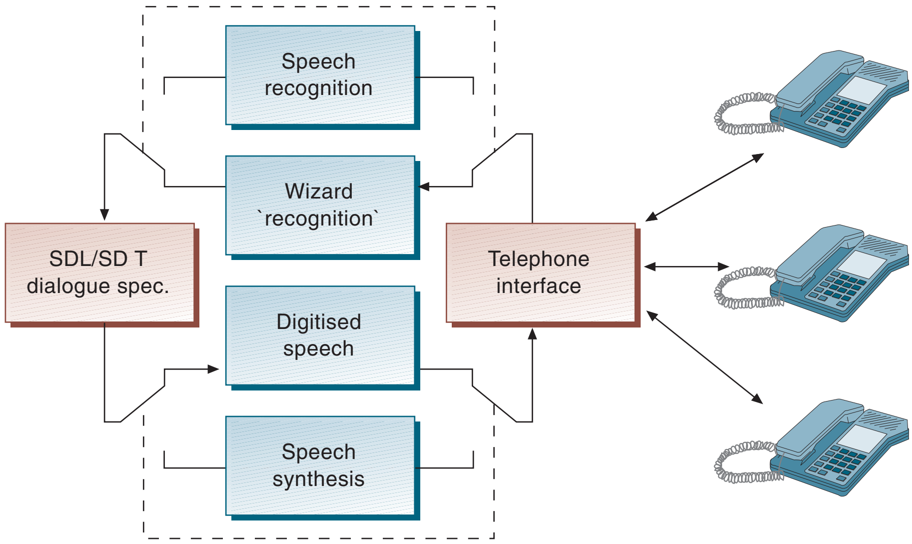


Figure 4 The test environment includes a tool where dialogues may be specified in SDL (Specification and Description Language). This tool contains a runtime module that allows specifications to be compiled and run in real-time in the same environment. This allows rapid prototyping and high consistency from specification to implementation (after [13])

2.4 Dialogue design for automatic speech recognition in number paging

To send a message (number) using an automatic numeric paging service requires tone signalling (DTMF) and telephones with * and # keys, but if this is not available an operator-based system is required. Operator-based numeric paging incurs high costs on the service provider. The approach to design an automatic number paging service for users with no access to DTMF, to reduce the use of operator-based services, is described by Hamnes, Amdal, Kulseth & Rugelbak [13] from Telenor Research and Development, Kjeller, Norway. The work was carried out in a speech-based dialogue experimentation environment, allowing easy switching between speech synthesis and digitised speech and between speech recognition and the *Wizard of Oz* simulation of recognition (see Figure 4). Traditional tools and methods, such as *Wizard of Oz* simulation, SDL (Specification and Description Language) and MUSE (Method for Usability Engineering) were utilised (*Wizard of Oz* = covert simulation of speech recognition and synthetic speech).

Hamnes *et al.* first set up some user requirements and design guidelines:

- The successful transaction of sending a number to a pager should not take noticeably longer to complete than the same transaction using an operator-based service.
- The user should not require any training other than a short instruction printed in the telephone directory (equivalent to the current automatic or operator-based services), and the dialogue should be easy to learn.
- 'User friendliness' should be equivalent to or better than the operator-based service.
- The service should issue an initial 'welcome' message, telling the user which service has been reached.

- The service should clearly differentiate between the pager number and the message number being sent to the pager.
- The service should give feedback of both the pager number and the message. At some point the whole number (pager or message) should be read back to the user, so that the user may either confirm or cancel if the number has been recognised incorrectly.
- The service should indicate to the user that the message has been sent and give a clear exit message (observing Ciganek's tenet of closedness, i.e. that a transaction should have a beginning, a task stage, and a defined end).
- The dialogue must be able to handle error situations, e.g. the user giving no input or the wrong input.

Hamnes *et al.* identified and analysed some dialogue parameters such as:

- Speech recognition technologies
- Grouping of digits for feedback
- Prompts
- 'Barge-in'
- Dialogue 'style'.

On the basis of this, four dialogue profiles were chosen for prototyping by assigning values to the different dialogue parameters (see Table 1). A *Wizard of Oz* simulation study was then carried out to evaluate the four dialogues with respect to the user requirements, and to compare them with the existing operator-based service.

The findings showed that only dialogue C (connected digit recognition, terse style) had a shorter completion time than the operator-based service. Dialogues A and B both had slightly longer completion times. This had been expected since they both apply isolated digit recognition, which leads to slower number input. Dialogue D (connected digit recognition, user-specified message length and a verbose style) gave a significantly longer completion time. This was partly due to subjects not using the 'barge-in' facility, which would have reduced the task completion time.

All four dialogues showed clear learning effects. Learning time was determined by comparing the actual time used with a reference time measured for each of the four dialogue styles. For dialogues A, B and C, the task completion time was very close to the reference times by the time of call number three. The task completion time with dialogue D was still significantly longer than the reference time by the third call, which may be due to callers not using the 'barge-in' facility.

All four automatic dialogues were rated higher than the existing operator-based service, particularly with respect to quality of feedback. Dialogues A, B and C were rated higher with respect to overall 'user friendliness'. Users preferred immediate feedback on each digit rather than feedback to groups of digits. The operator-based service, however, was rated higher with respect to subjective feeling of time used, with the exception of dialogue C, which used the most sophisticated recognition technology.

In conclusion, Hamnes *et al.* observe that by systematic approach based on the MUSE method, application of guidelines and analysis of dialogue parameters, many design alternatives could be ruled out at an early stage of the design process. The analysis supported hypothesising about the usability of certain combinations of dialogue parameter values. Due to low number of subjects, the validity of their quantitative results may be low, but findings were fairly consistent and supported hypotheses made prior to the study. 'Wizard errors' (i.e. subjects detecting the deception) were low and their effect should be minimal.

2.5 Performance, preference and design of speech recognition systems

Even though the technology now exists for implementing applications using speech recognition, we should not assume that speech recognition interfaces necessarily would be preferred over touch-tone or other interface types by all users. In a previous study, Fay [14] had found that users preferred touch-tone (DTMF) in a college registration service, but speech recognition in a telephone disconnect service.

D. Karis of GTE Laboratories, USA, studied two different services to get data on user preferences of speech recognition vs. touch-tone, user ratings and acceptance of a system as a function of the recognition performance experienced, and general and specific user interface design techniques [15]. He also addressed the design of speech-only interfaces as well as dual-mode (speech & touch-tone input) interfaces.

Data was gathered while developing two GTE services: *Speech Connect*®, an automated business directory service where callers can speak the name of the business they want to reach, and *InConnect*®, a call management service. In

Table 1 Dialogue profiles

Profile name	ASR techn.	Grouping of pager #	Establishing msg length	Grouping of msg	Digit prompt	Barge-in	Dialoge style
A	isolated	3-2-3	'end'-command	none	beep	no	terse
B	isolated	single digit	user-spec. length	single digit	feedback on digit	yes	verbose
C	connected	whole number	'end'-command	none	none	yes	terse
D	connected	3-2-3	user-spec. length	automatic	none	yes	verbose

Speech Connect® the active vocabulary is over a thousand words at some points, the system is speaker independent, does not support 'barge-in' and there is no touch-tone option. *InConnect*®, on the other hand, has a small vocabulary, uses both speaker-dependent and speaker-independent recognition, permits 'barge-in' and includes a touch-tone option that can be used instead of speech recognition. Some traditional methods were used, such as scenario-based usability evaluation, where subjects were given representative scenarios to follow, other laboratory evaluations, where subjects performed each task twice, once using speech and once using touch-tone, and large-scale data-collection efforts.

Karis found that although the two services differed on various aspects, several similar user-interface strategies were used successfully in both cases. It was found effective to provide an initial brief prompt followed, if the caller did not respond, by more detailed prompts. Effective error recovery was essential. Options include having the caller spell out the word, reverting to discrete vs. continuous speech input, providing more detailed instructions, or modifying the dialogue to improve recognition accuracy. For example, after two failures of *Speech Connect*® to recognise the name spoken by a caller, the caller is first asked to choose a business type before speaking a particular company name.

Karis also attempted to track callers' acceptance of each service as a function of speech recognition performance, and to determine how low the speech recognition can be and still result in a successful interaction and positive evaluation of the service. Karis found that the acceptable recognition level depends strongly on the efficiency and success rate of the error recovery routines. Services can be successful with recognition accuracy significantly below 90 %, provided the transaction has a high success rate and interaction is smooth and efficient. In the *InConnect*® service, which provided a touch-tone alternative, the majority of subjects in the study preferred touch-tone, which was faster than speech. A large minority, however, preferred speech because they did not have to remember key mappings. An interesting difference was found between college students and older adults, with college students showing a strong preference for touch-tone and older adults preferring speech recognition.

In conclusion, Karis proposes the following five specific design rules:

1. Provide brief initial prompts followed, if necessary, by more detailed prompts.
2. Reduce prompts when they occur multiple times in a single session (called 'tapering').
3. Do not repeat the same prompt verbatim over and over (a change in prompt indicates that the system is aware that the same event is recurring).
4. Provide additional information after a problem (this is an important component of effective error recovery).
5. Provide mechanisms to give the user control (e.g. to move from one part of the interface to another, or to get help when needed).

2.6 Fuzzy reasoning in computer aided telephony

For inquirers who do not know exactly what information they seek or have no precise idea about which questions to ask, there was until now no other option than talking to a person with a general knowledge of the field. Is it possible to design an automated computer based telephone system that can answer imprecise questions?

Zajicek, Brownsey, Palau & Lippmann, the School of Computing and Mathematical Sciences, Oxford Brookes University, UK, describe the development of a computer based telephone answering system designed to answer initial calls to the university concerning courses on offer for associate students. The central problem was uncertainty. The system conducts a question and answer type dialogue to establish which areas the callers might be interested in and give them a short list of suitable course module names. The system uses fuzzy logic [16,17] to model the uncertain reasoning in the dialogue and the vagueness within the database itself [18].

Up to now, information and a person, who enters into a dialogue with the information seeker, has provided advice. The uncertainty is resolved, at least in part, in the dialogue process. An example of a typical dialogue between a caller (C) and provider (P) is given in the human-to-human dialogue fragment cited below:

C: "I was thinking of doing something with psychology."

P: "Have you had any previous experience of work in this area?"

C: "No, but I have always wanted to work with people."

P: "Do you already have any relevant qualifications?"

C: "Not in anything to do with psychology."

P: "Have you thought of something to do with caring?"

C: "No, but I'd like to know more about that. Could you tell me something, please?"

P: "Well, it covers a number of different vocations. For example, would you be interested in mental health nursing?"

C: "Quite possibly – I had never thought about it before to be honest."

This dialogue fragment shows different kinds of uncertainty. The caller is vague about his goals. Although the provider knows that about 65 % of all calls are to enquire about psychology and a hint in this direction may be a hit, the provider is still vague about what level or kind of vocation to suggest, but a specific question elicits the fuzzy reply "Quite possibly".

Up to now, computerised telephone answering systems have been using hierarchical structures, i.e. branching tree-structures leading from more general to more specific questions giving a 'hard-wired' dialogue even though some systems may allow backtracking. Zajicek *et al.* wanted the dialogue to retain some key features of the human-to-human dialogue. A good information provider should be discursive with an inquirer who has unclear goals. Suggestions should reflect what has happened in the dialogue up to that point and previous experience with enquirers.

The authors then deal with knowledge representation and user modelling. Domain knowledge can be represented as various elementary groupings containing two 'collective terms' and one 'atomic term'. An elementary grouping would contain something officially recognised such as 'French studies'. A collective term would contain more informal expressions used by the public such as 'twentieth century culture', 'academic subjects' or 'science interests'. An atomic term would contain names of individual course modules, such as *Mod-*

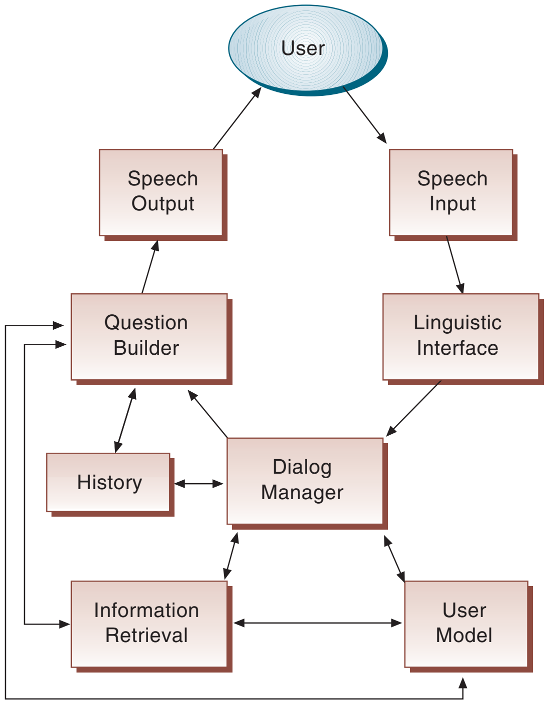


Figure 5 The architecture of the telephone answering system showing the various components arranged around the central dialogue manager (after [18])

ern French Literature, Quantum Physics, Business Ethics or *Introduction to Spreadsheets*. For knowledge representation the authors used a simple fuzzy database consisting of a single relation **R**. The user interest model consisted of a set of pairs created by assigning a user interest value from a linguistic scale for each term in the knowledge domain. These values of the scale were:

- Not interested
- Not very interested
- Undecided
- Quite interested
- Interested.

The user interest values would be updated as a result of the dialogue. Initially, all values were set to 'Undecided'.

The system architecture of the eight components is shown in Figure 5. The *User* accesses the system by typing in questions to the *Speech Input* (in the future speech recognition is envisaged). In the *Linguistic Interface* the input is transformed into objects usable by the deeper layers of the system. The *Dialog Manager* starts off the dialogue, infers what to ask about next, decides from the responses what updating is required, chooses what information to offer, and decides when to stop the dialogue. The *User Model* holds the latest version of the user interest model and updates it as necessary from input from the *Dialog Manager*. The *Information Retrieval* is a simple database query and answer system. Its role is to determine whether the model is sufficiently refined to offer advice or if the end of the dialogue has been reached. The *Question Builder* constructs questions aimed at eliciting useful responses from the User about the terms introduced, by providing context to the questions and a variety of question forms. The *History* unit simply keeps track of all the (essential items) of questions and answers. And, finally, the *Speech Output* is simply a command line displaying the question constructed by the *Question Builder* (in the future synthetic speech is envisaged).

2.7 Assessing quality of synthetic speech

The International Telecommunication Union (ITU) has investigated subjective methods for assessing synthetic speech quality [19,20,21]. According to ITU-T Recommendation P.85, the references for subjective testing of speech quality include natural speech, modulated noise reference (MNR) and time-frequency warped (TFW) speech. Although such references are important for obtaining stable assessments, Rec. P.85 does not clearly state which reference should be used or how TFW speech should be generated.

Toshiro Watanabe, NTT, Japan, proposes a set of speech samples generated by time frequency warping based on phase modulation (TFW-PM) from low bit-rate sampled speech [22]. He suggests that speech quality depends on three variable parameters:

1. Number of PARCOR synthesis coefficients (p)
2. Frame length (L)
3. Maximum phase shift ($\Delta\theta$).

In order to determine one parameter variable and optimise the other parameter values, he performed a subjective experiment using TFW-PM references generated by combining those parameters and four kinds of synthetic speech based on the Rec. P.85 assessment method. Each test sample, spoken in Japanese, was assessed on five qualities:

1. Overall impression
2. Acceptance
3. Articulation
4. Pronunciation
5. Voice.

He found that the $\Delta\theta$-parameter had the greatest influence on the opinion scores for over-all impression and decided that this parameter was a variable of TFW-PM and that qualities of synthetic speech were expressed in equivalent $\Delta\theta$. Two other experiments were performed under quiet and noisy listening conditions using both MNR and TFW-PM references in a similar manner to the first experiment. His results show that synthetic speech qualities in terms of equivalent $\Delta\theta$ are more stable than equivalent Q (dB) with both trained and naïve subjects.

2.8 Voice reaction time in interactive services

Transition from automatic interactive telephone services where users respond by key presses to interfaces where users respond by speech raises the question if the psychological mechanisms involved in the two ways choosing and responding are comparable. When making choices, these can be signalled by speech or pressing keys. In a classical experiment Hick [23] found that reaction time increases logarithmically with the number of stimulus-response choice alternatives; universally known as *Hick's Law*.

With interactive telephone services, choices are presented audibly rather than visually. This corresponds to a situation described by Davies *et al.* [24] who experimented with auditory stimuli and voice responses (described as 'compatible'), while responses in traditional interactive telephone services are manual (i.e. 'incompatible').

Although responding vocally to auditory stimuli may appear 'natural', it does not necessarily mean that it is 'compatible' (as defined by Davies *et al.*). Since the ear and vocal systems utilise different parts of the central neural system, there are no direct neural links between them. Therefore, any strong relationship that may exist is the result of over-learning.

Opposite to motor responses, comparably little is known about vocal responses. Thomas [25] from Philips Corporate Design, Netherlands wished to determine whether vocal reaction times follow similar patterns as motor reaction times. He replicated an experiment by Mowbray [26] where subjects had to name numerals from a familiar response set (their native language) and from an unfamiliar set (a strange language). His hypothesis was that vocal choice reaction times for the familiar response set (*in casu* English numerals) would remain constant over number of choices, while the choice reaction times for the unfamiliar set (Hebrew numerals) would show a logarithmic increase consistent with *Hick's Law*.

The results (see Figure 6) clearly support this hypothesis. With unfamiliar (Hebrew) material the mean response times grew logarithmically over sets of two, four or eight digit choices, confirming *Hick's Law*, whereas with familiar (English) material the mean response time for two, four and eight choices remains constant, which is consistent with Mowbray's finding that no appreciable differences exist when the stimulus-response material is over-learned.

Thomas deduces from these findings that responding vocally, rather than manually, should have similar information processing requirements. This conclusion, however, remains to be tested. It is questionable whether just replacing key presses with vocalisations will be sufficient to achieve an acceptable speech interface. He maintains, however, that with increased familiarity with the response material and as speech technology matures, indicating choices by speech will become as intuitive as pressing keys.

2.9 Evaluating synthetic speech

Today few interactive touch-tone telephone services employ synthetic speech, i.e. voice produced by a TTS (= Text-to-speech) system where input of a string of text characters is converted into output of speech waveforms. This is mainly due to the fact that synthetic speech is still of much lower quality than recorded natural speech. The main problems are [27]:

- Intelligibility
- Comprehension
- Listening fatigue
- Increased attention
- Cognitive load.

Cristina Delogu and Andrea Paoloni from Fondazione Ugo Bordoni in Italy have addressed the task of evaluating speech synthesis systems for telecommunication services [28].

They find that the nature of the speech material, as well as the specific tasks given, both have implications for evaluating synthetic speech. The authors review some traditional tasks used for evaluating speech quality:

- Transcription (writing down what is heard)
- Preference judgement (giving relative preference between alternatives)
- Scale judgement (indicating preference on Lickert-scales, e.g. 1 – 7)
- Comprehension (answering multiple-choice questions about material).

The authors claim that experimental findings from laboratory tests cannot make reliable predictions about TTS synthesis performance in real life applications since these do not take into account significant cognitive factors involved in the perception of the synthetic voice. They also speculate about merits of alternative new methods, such as measuring the cognitive load in perceiving a message, measuring reaction time, i.e. the time to write down the message in transcription tasks, and concurrent tasks, i.e. performing a second task while listening to the message.

The authors present the results of several laboratory experiments where listening difficulty, comprehension, cognitive load and listening attention to synthetic speech (Italian) was measured. These show that target words embedded in sentences spoken by a clear natural voice and by the best *vocoders* were recognised faster than words in sentences uttered by format-based TTS systems. However, when re-tested a week later the voices from the TTS systems showed a larger decrease in reaction times than the natural voice and the good *vocoder,* suggesting that performance in voice recognition will increase with familiarity with synthetic speech.

Speech comprehension tests with a natural voice and a TTS synthetic voice, with multiple choice question responses, did not show any significant differences in the answers between the two conditions. The authors speculate that the subjects may have used different strategies, involving linguistic knowledge. This could have cancelled any effects of the unlike speech qualities of the two voices. To test this hypothesis the authors performed an experiment where the reaction times to recognise computer generated clicks while listening to texts read by natural and TTS-generated voices was measured. The results showed that reaction times for clicks embedded in synthetic speech were longer than in natural speech, while no difference in comprehension was found. This suggests that synthetic speech leads to a greater workload on the listener.

In an attention test, fluctuations in vigilance (*viz.* to indicate 150 clicks) were measured for 22.5 minutes long texts read by natural and synthetic voices. The subjects also had to answer questions about the texts. The results showed that subjects listening to a human voice reveal cyclic fluctuations of attention of about 3.5 minutes between 'power peaks', while those listening to a synthetic voice had cycles of about 2 minutes. Thus, subjects show a more rapid fluctuation of attention when listening to

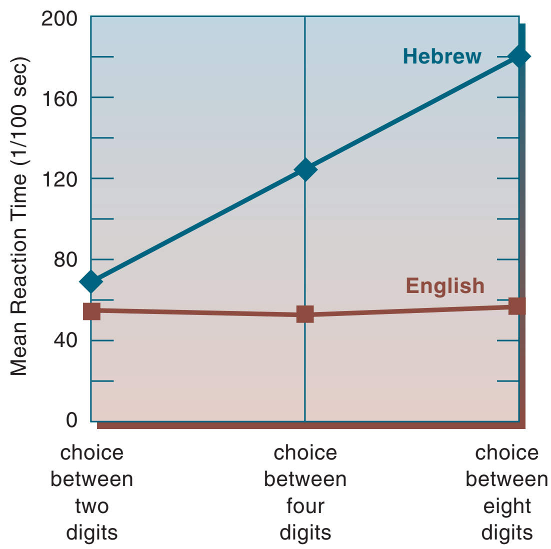


Figure 6 Mean choice reaction times in 1/100 sec. (ordinate) plotted as function of number of digits to choose between (2, 4 or 8, abscissa) with responses in English (■) and Hebrew (◆) as curve parameters. Hebrew responses, which are still in the process of being learned, conform to Hick's Law, *while English responses, which are well known, do not (after [25])*

a synthetic voice than when listening to a natural voice, but exhibit no differences in comprehension of the text passages.

Delogu and Paoloni also describe an evaluation of an Italian automated 'reverse' telephone directory assistance service (ARDA). The user enters a telephone number via the keypad and the TTS system retrieves the listing information from a database and presents it over the telephone. Greetings and information about the service are given by pre-recorded natural speech, while surnames, proper names, street addresses and names of cities are presented by TTS synthetic speech. This service was compared to a simulated service using a natural voice. The subjects were to write the surnames and addresses associated with 20 dialled numbers.

Their results show that natural speech gave 20 % more correct surnames and addresses than synthetic speech. This is not very surprising since surnames, and to some extent street and place names, do not always follow orthographic rules. Also, there are no rules for where to put the stresses when pronouncing surnames and here synthetic speech systems are more at a disadvantage than a human speaker is. Surnames tend to be isolated words, while addresses have stronger linguistic, geographical and historic context, which may account for differences in intelligibility. No differences were found in task completion between the two groups. In general, the users' acceptance of the TTS service was low and overall impression was poor: the subjects reported that the sound was not clear, pronunciation was annoying and the voice was unpleasant.

In conclusion, Delogu and Paoloni observe that there were at least three important differences between the laboratory conditions and the field test conditions which may all have large influence on the results: *ambient noise* (the laboratory was at least 20 dB quieter than an office and 40 dB quieter than a kiosk in the street); *prosody* (surnames need correct pronunciation to be recognised) and *task complexity* (in the laboratory one name at a time was presented, while in the field the next piece of information was given by the system while the subject was still writing the previous information, thus creating interference). They end by asserting that more research should be directed towards the underlying cognitive processes that take place when listening to speech to interpret the evaluation results of TTS systems and make informed decisions on future applications of synthetic speech.

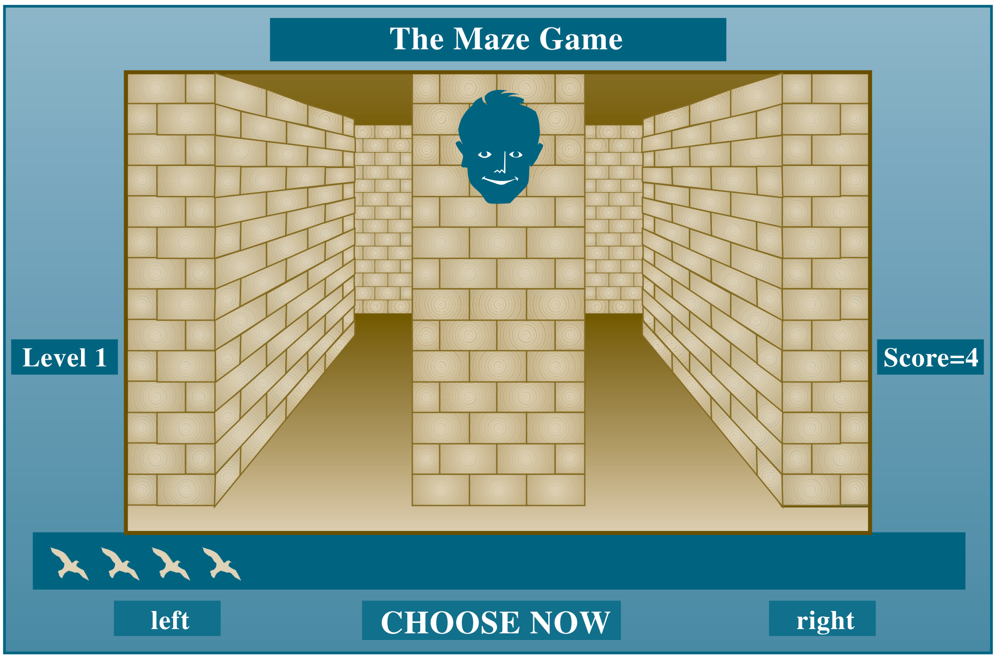


Figure 7 An imitation of the Maze Game screen display showing a talking head and a prompt at the bottom of the screen to choose the left or right side (after [30])

3 Video communication and multimedia

Interpersonal video communication, e.g. video conferencing, video telephony and now video on the PC and via the Internet, has become possible with the availability of wider bandwidths and efficient picture coding algorithms, and is now being taken up in an in creasing number of applications. This has raised a host of new human factors issues, some of them are dealt with in a companion article in this issue of *Telektronikk (Human Factors in Videotelephony)*. Other human factors issues dealt with here include such areas as, trust in talking heads, evaluation of subjective picture quality, use of video-communication for deaf people using sign language, avoiding the 'puppet theatre' effect in video displays, 3D pointing devices in stereoscopic displays and sign language interpreting service on broadcast digital TV.

3.1 User trust in 'Talking Heads'

Walker, Sproull & Subramani [29] identified three important issues about the use of 'simulated personae' (*talking heads*) in human-computer interaction: 1. Are people actually willing to interact with a talking head; 2. Will people be so distracted by a talking head that it will interfere with task performance; 3. How 'human' is a talking head and will it evoke any 'social response'.

To address these questions, Taylor, McInnes, Foster & Jack [30] from Centre for Computation Interface Research, University of Edinburgh, performed an experiment to assess user trust in talking heads as information providers. The experiment scenario was a video screen-based computer game where the players were to walk through an underground maze of rooms and passages and collect silver coins (value, 1 point) and gold coins (value, 100 points). Three different advisory agents appear, each in a separate game, and are compared for user trust (a control game with no advisory agent was also included). At each turn in the game the players found themselves facing two passages, one to the left and one to the right (see Figure 7). To collect the silver coins hidden in one of the two passages the players were prompted at the bottom of the screen to choose 'left' or 'right'. On some occasions a gold coin was hidden and a speaking 'agent' advised players to choose 'left' or 'right' (except for the no agent control condition). Three different 'agents' were used:

- Talking head agent (synthetic video)
- Disembodied voice agent (voice only)
- Real video head agent (video of real person).

At the start of the game, players always found the silver coins by choosing consistently the same side for each new move. When the agent appeared, it gave wrong advice on where to find the gold coin (always the opposite side of the silver coins), i.e. if the player followed the advice no gold coin was found, but neither was a gold coin found if the player ignored the advice. After some more silver coin moves where players always found coins on the same side as before, the agent appeared once more to give advice on which side to find the gold coin.

For players who had followed the agent's advice on its prior appearance the advice was now pre-arranged to be correct. If players again followed the agent's advice (in the face of contrary experience of misplaced trust) they were rewarded with a gold coin. Players who did not follow the advice did not receive any coin.

For players who had *not* followed the agent's advice on the first occasion, the advice was now pre-arranged to be wrong. If they followed the advice they were not rewarded, but if they again rejected the agent's advice they received a gold coin. The game was thus designed to ensure that each player was given correct advice once and wrong advice once to counter any bias either against or in favour of taking advice from an agent on subsequent interactions. The net result of the design was that players who were consistent in their decision to either accept or reject the agent's advice were rewarded with a total of one gold coin (in addition to the silver coins collected).

The players each completed four games (one for each condition) and within each game made four decisions regarding the gold coin. These decisions are the main focus of the experiment. For each agent type, the number of decisions that departed from the previous experience (i.e. the side silver coins were found) was tabulated and the means computed. The percent deviation for each condition were:

- Talking head, 48 %
- Disembodied voice, 46 %
- Real video head, 45 %
- Control condition, 37 %.

The control condition, thus, gave the least number of departures, while the differences between the three agents are small. The difference between the control condition and the *mean* of the three agent conditions was significant ($p = 0.041$) as were the differences between the control condition and each of the three agent conditions.

The time to make the decisions were measured and averaged, and the ratio of the average decision time for the first gold coin and the mean decision time for the silver coins was computed for each condition. The control condition yielded significantly shorter decision time ratios than any of the three agents.

Taylor *et al.* conclude that advice *does* make a difference: there were more departures from immediate prior experiences in the advice conditions than in the control condition, and the players who were advised took longer to reach their decisions.

With respect to the decision on the first gold coin in each game, there were moderately significant differences between the talking head on one side and the disembodied voice and real video head on the other side. It appears that the talking head was more persuasive than either the real video head or the disembodied voice. The authors speculate that the differences found between the talking head and the other conditions may suggest that a sufficiently realistic human-like synthetic 'talking head' might be trusted more.

3.2 Subjective evaluation of audio-visual quality

Perception of audio-visual quality of multimedia services, including real-time interpersonal communication and interactive remote data retrieval services, may not only depend on the quality of each of its two principal components (audio and video) but also on their interactive (synergetic) quality. McGurk & MacDonald [31] and Sekiyama & Tohkura [32] demonstrated perceptual fusion of auditory and visual (lip-reading) information *(McGurk effect)* in speech recognition under audio-visual discrepancy (audio and video out of sync.).

Contin[*], Quacchia[*], Galati[#], Miceli[#] & Turi[#] [33] from [*]Telecom Italia, CSELT, Turin, Italy, and [#]University of Torino, Italy, have pursued this basic idea in a series of experiments to find a method for subjective evaluation of audio-visual quality. The test material consisted of six video excerpts with audio, each lasting for 15 seconds. These excerpts were compressed by a commercial MPEG-2 *codec* and recorded in various combinations source quality and of coded audio and video bit-rates (256 kbit/s, 128 kbit/s and 64 kbit/s for audio, and 12 Mbit/s, 8 Mbit/s and 4 Mbit/s for video).

Pairs of sequences of the same material were presented according to the ITU-R Double Stimulus Continuous Quality Scale (DSCQS) protocol. One sequence in each pair was always uncompressed (source). The subjects were requested to evaluate each sequence on three quality parameters; 1) 'colour reproduction', 2) 'smoothness of movement,' and 3) 'picture sharpness', using the continuous ITU-R quality scale for the DSCQS method. The reproduction of 'low' and 'high' audio frequencies was also assessed. In addition to the audio and video evaluations, an evaluation of the 'overall audio-visual quality' was requested.

The results of Contin *et al.* [33] show that for any video bit-rate, high frequency audio quality increases with increasing audio bit-rates, and for any audio bit-rate, the subjective 'picture sharpness' increases with increasing video bit-rates. This suggests that both audio and video bit-rates significantly influence the perceived quality of high frequency sounds. Analysis of variance also showed that the 'video bit-rate' and 'sequence' had significant effects on the 'smoothness of movement' quality, and that the 'audio bit-rate' had a small, but significant, effect on the quality of 'colour reproduction'.

For evaluation of low frequency audio quality, only the 'audio bit-rate' and 'video sequence' factors were significant. Surprisingly, the 'video bit-rate' was not significant for the 'overall audio-visual quality'. The authors warn, however, that one should be careful about generalising that the 'video bit-rate' is an insignificant factor for 'audio-visual quality'. They conclude that in order to avoid any evaluation bias, 'overall audio-visual quality' should be assessed in separate test sessions.

3.3 Coding degradation and recognition of emotion

In remote interpersonal video communication, such as video-telephony and videoconferencing, the visual feedback of the gestures and facial expressions of the interlocutors is the *raison d'être* for the medium. However, the recognition of non-verbal information, attitudes and emotions may be lessened by the picture degradation introduced by the video coding algorithms.

Contin*, Quacchia*, Galati#, Miceli# & Torrini# [34] from *Telecom Italia, CSELT, Turin, Italy, and #University of Torino, Italy, carried out an experiment to identify the thresholds for correct identification of human emotions at very low bit-rate coded video sequences. Twelve silent video sequences, showing children expressing six different emotions (but not showing any context), were compressed by the MPEG-4 video compression algorithm and up-sampled to the Common Intermediate Format (CIF) (360 pixels, 288 lines, 25 Hz, 4:2:2). Three bit-rates, 16 kbit/s, 24 kbit/s, 48 kbit/s, and the uncompressed CIF sequences were used for each video sequence. Subjects had to indicate which of the following emotions the child in each sequence expressed:

1. Anger
2. Joy
3. Disgust
4. Surprise
5. Sadness
6. Fear
7. Other.

Subjects also had to rate the intensity of the emotion shown and the degradation of each video sequence on two Lickert scales.

In general, the results show that lower compression bit-rates lead to lower correct recognition of emotion, but not to the same extent for all emotions. And even in the non-compressed versions of the sequences used, correct recognition depends on the emotion (ranging from 63 % for 'surprise' to 92 % for 'joy'). Perception of the intensity of the emotion, however, does not seem to be affected by the compression. As expected, the subjective picture quality decreases with increasing compression ratio up to 20 % for the lowest bit-rate. The authors conclude that further studies are required to determine what factors are important for remote interpersonal video communication at low bit-rates in e.g. mobile networks.

3.4 Sign language and 64 kbit/s picture quality

Several studies have shown that the 64 kbit/s ISDN video telephone is sufficient for remote interpersonal sign language communication by deaf people (see e.g. [35]). However, it has also been shown that remote sign language conversation by video telephone is more stressful than real face-to-face conversation [36,37,38].

Kamata, Yamashita & Hirama [39] of Utsunomya University, Japan, have studied the relationship between picture quality of 64 kbit/s ISDN video telephones and sign language conversation. Using Scheffe's paired comparison method they performed perception tests with deaf and hearing subjects and different picture quality.

The results show that the deaf subject preferred pictures with higher *temporal* resolution (faster frame rates) to pictures with higher *spatial* resolution (more detail), while the hearing subjects preferred higher spatial resolution to higher temporal resolution. These tests did not actually use real sign language conversation, but a simulation, and only one deaf subject, but the authors are confident that their findings are representative for deaf people using sign language. The authors suggest that the party receiving the video picture should be able to control the video coding parameters of the sending party to suit their special requirements.

3.5 Sign language in digital video broadcasting

Real-time sign language interpretation in a small window in the broadcast picture of TV programmes has been provided for the deaf in Norway on special occasions, e.g. political debates in connection with general elections. The Norwegian Broadcasting Corporation (NRK) has for several years investigated the possibility of offering sign language video as a permanent service. With the advent of the digital video broadcasting (DVB) standard, a co-operation between NRK and Tandberg Television, was initiated to study the technical possibilities offered by the standard for sign language video. The goal was to develop a permanent additional sign language transmission in parallel with the ordinary TV programme as an optional feature of the TV receiver.

Fosse & Dolvik [40] of NRK and Tandberg Television, respectively, describe a project to develop such a service. With the imminent adoption of the DVB standard it was deemed irrelevant to consider transmission of the service via the NICAM multiplex.

The authors give a brief technical description of the basic MPEG-2 DVB system for transmitting a user selectable service where sign language is transmitted in parallel with the ordinary TV programme and shown as a small picture in the upper left hand corner of the TV screen. Thus, hearing people will not have their viewing disturbed by the sign language.

Two prototype satellite receivers were developed to demonstrate the system. The system has been demonstrated on several occasions. Its general configuration is shown in Figure 8. The pilot project has proven the practicality of the proposed system. Work will now concentrate on implementing the system and find industrial partners who will put the necessary consumer equipment on the market.

3.6 'Informal' video-communication

Informal communication designates those spontaneous, non-mediated, informal face-to-face exchanges, usually between two or three people, that occur in the work place in chance meetings around the office, during coffee breaks, in the tea-room, or in informal gatherings at conferences, and that fulfils important functions such as information exchange, trying out ideas, harmonisation of attitudes, etc. Most non-mediated, informal communication is characterised by being:

- Unscheduled
- Confidential
- In real time
- Interactive
- No agenda
- No moderator.

To derive user requirements for telecommunication systems that support informal communication, Suwita and Mühlbach [41] from Heinrich-Hertz-Institut für Nachrichtentechnik, Berlin, studied the nature of informal communication by interviewing ten male and nine female employees from different work areas such as, court of justice, school, research institution, hospital, industry, etc. In the semi-structured interview, some questions centred about occasions, locations, durations, benefits and disadvantages of informal communication. There was one question whether the interviewee could imagine informal communication by means of video communication.

The study confirmed the authors' belief that informal communication was considered important, even indispensable for work. Besides improving the social work climate and the subjective well-being of the employees in general, various positive aspects of informal communication were referred to. These include bringing team members up-to-date, that it stimulates new ideas and thoughts and that it helps to solve every-day work problems in an efficient manner. Negative effects were also mentioned; e.g. that people may be disturbed in their work and that it may upset the social climate by gossip and rumours.

Among the most crucial conditions for informal communication to take place are confidentiality and privacy; people sometimes make statements and express opinions during informal chats that they would rather not make in a formal meet-

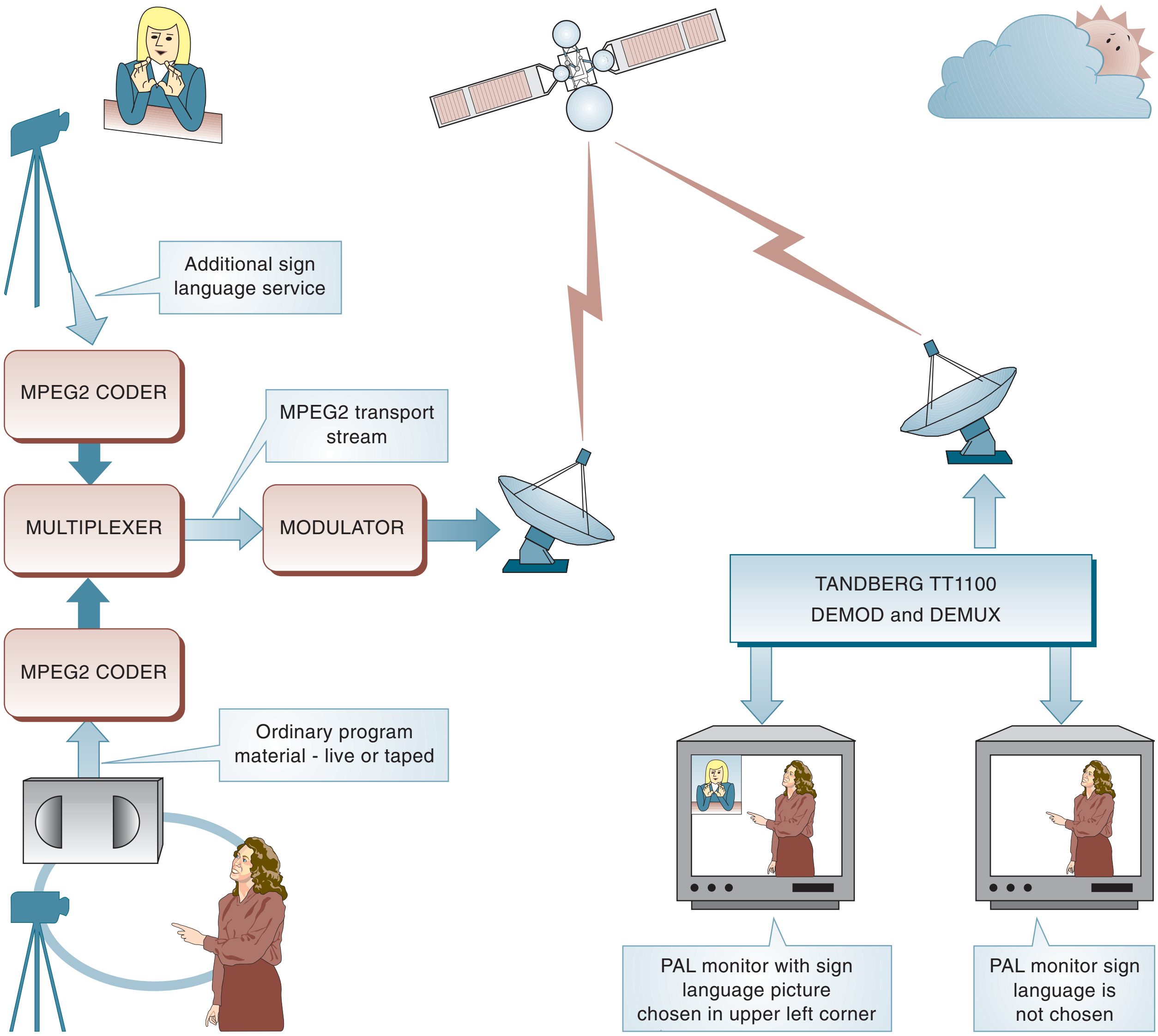


Figure 8 The Digital Video Broadcasting (DVB) program chain showing the addition of a real-time sign language interpretation service signal to the main programme material, and the viewer's option to select or reject the live sign language interpretation service provided in the upper left corner of the TV screen (after [40])

ing. Environmental factors that support informal communication are spatial proximity, availability of common rooms and open office doors. Informal encounters last only a few minutes, seldom longer than 15–20 minutes, but may recur during the working day.

On the question of video-communication, the majority of interviewees could hardly imagine to have informal confidential chats by means of video-communication. They felt that it did not provide enough social presence and expressed grave concerns that video-communication does not offer sufficient privacy and confidentiality, e.g. fear of being observed (even videotaped) or having no control of who could watch them.

The authors conclude that informal communication is very important in the work context, thus designers of telecommunication systems for tele-co-operation and distance work should take this into consideration. The authors also expect that systems that support informal communication will enhance user satisfaction and improve work efficiency.

3.7 The 'puppet-theatre' effect in 3D desktop displays

In the common 3D desktop displays, only a limited field of depth around the screen surface can be utilised without causing eyestrain. This leads to people and faces being reproduced at an unnaturally small scale, which causes a 'puppet-theatre' effect (see Figure 9), preventing a high degree of presence in interpersonal communication applications.

Hopf & Runde [42] of the Heinrich-Hertz-Institut in Berlin describe a way to present scenes at some distance 'behind' the screen (virtual image) to avoid the miniaturisation effect in an autostereoscopic (3D) desktop display for communication applications. They report on a study where they shifted the plane of accommodation with a collimating optical system involving Fresnel lenses and spherical mirrors (see Figure 10).

Prussog *et al.* [43] had found that participants in stereoscopic videoconferences felt uncomfortable wearing 3D goggles or other head-mounted equipment. The authors therefore conclude that 3D displays should be *autostereoscopic,* i.e. that the image can be viewed in 3D without any supplementary devices like shutters, polarising glasses or similar gadgets, while allowing the viewer some freedom of movement. Their solution is to place an optical collimating system

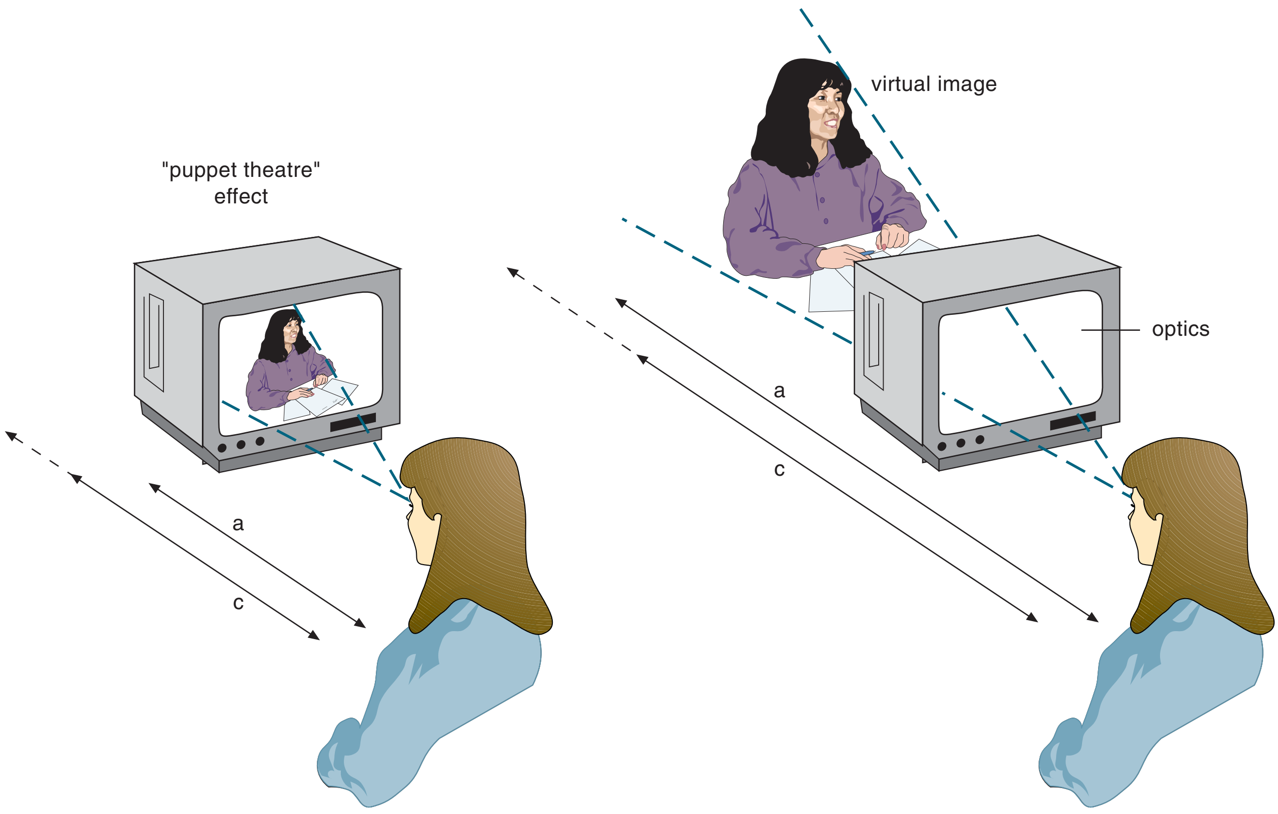


Figure 9 The principle for the adaptation of accommodation distance by means of optics to defeat the 'puppet theatre effect' (a = accommodation distance, c = convergence distance). The small image on a screen can be experienced as a life size 'virtual image' if the accommodation distance of the visual system can be tricked into perceiving the image as being further away than it actually is (after [42])

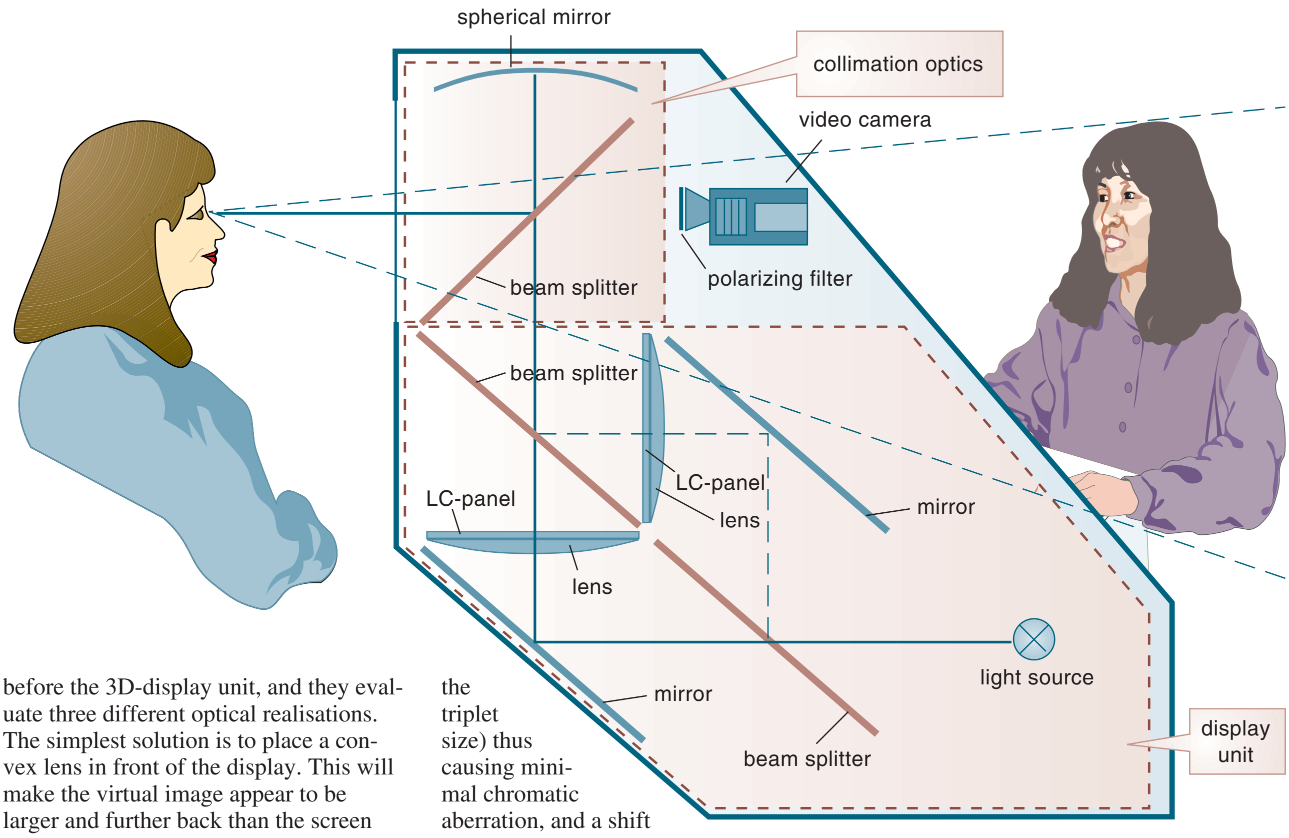


Figure 10 A 3D desktop display unit for multimedia. The collimation optics (upper part) will project a virtual 3D image from the display unit (lower part) so as to increase the perceived accommodation distance and defeat the 'puppet theatre effect' (after [42])

before the 3D-display unit, and they evaluate three different optical realisations. The simplest solution is to place a convex lens in front of the display. This will make the virtual image appear to be larger and further back than the screen surface. It is important to find the right parameters.

To avoid eyestrain, the accommodation plane should be placed some distance behind the screen so that differences between accommodation and convergence are minimised. This should result in an increase in the visible volume of depth, thus escaping the 'puppet theatre' effect. The other optical solutions evaluated were; a concave spherical mirror and beam splitter, and an off axis (45°) concave spherical mirror.

Hopf & Runde found that the use of a Fresnel lens gave unacceptably large chromatic aberrations even at low magnifications and on-axis eye-position (the difference in displacement of the red and blue components of white light was six times the pixel (triplet) size). The off-axis spherical mirror system reduced the spatial resolution, even at low magnifications, by enlarging the pixel (triplet) size over four times. With off-axis systems or high magnifications, deviation from correct focus leads to asymmetrical enlargement of the pixels (astigmatism).

The use of an on-axis concave spherical mirror and a beam splitter gave the smallest pixel magnification (only ± of the triplet size) thus causing minimal chromatic aberration, and a shift of eye-position of 150 mm produces an enlargement of only 1.8 of the pixel size. Some curvature of field makes it necessary to change accommodation for viewing different parts of the screen; the variation in focus amounts to ±150 mm. However, the beam splitter in the on-axis spherical mirror system causes a loss of about 75 % of the light. Geometric distortions arise in all three systems, but the best results were obtained with the on-axis concave spherical mirror; observing from an on-axis position leads to distortion of only about 0.6 %.

The authors conclude that the on-axis concave spherical mirror and beam splitter is the best system. It produces no chromatic aberration or distortion and resolution in on-axis viewing is acceptable; the loss of light, however, must be compensated for by increasing the display luminance.

3.8 2D-pointing in 3D space

Addressing objects in 3D space in interactive co-operative stereoscopic applications requires stereoscopic pointing. The most used pointing device is the standard computer mouse. But how are the 2D co-ordinates of the mouse translated into three co-ordinates in 3D space to create a user-friendly pointing device. (Using a mouse has the added advantage that ordinary 2D computer applications can be performed without changing pointing device.)

Runde & Suwita [44] from the Heinrich-Hertz-Institut, Berlin, have devised and evaluated three different 2D-to-3D pointing devices for use with interactive stereoscopic desktop displays. These are called the 'Surface pointer', the 'Floating pointer' and the 'Torch'. In stereoscopic images the pointer (3D cursor) has six degrees of freedom: three translations of movement in the three orthogonal dimensions and three rotational directions for changing orientation. The mouse, however, has only two degrees of freedom, movement along the two co-ordinates of the plane, plus one, two or three mouse buttons for clicking on various files and applications.

The ‘Surface pointer’ is a three-dimensional arrow that can be moved by the mouse in the two dimensions of the screen plane. The third (depth) dimension is automatically adapted so that the pointer always appears on the surface of the objects it touches. If the pointer is moved away from an object, the pointer does not jump to the background, but keeps to the same plane as the object. The pointer can be rotated in 30° steps in two planes by pressing the left and right mouse buttons, respectively. This pointing system relies on information about the 3D surfaces of displayed objects.

The ‘Floating pointer’ is of a three-dimensional arrow. Movement of the mouse determines the position of the 3D pointer in a plane parallel to the screen surface. Moving the mouse towards or away from the user while pressing the left mouse button changes the plane of depth in the 3D picture. Moving the mouse while pressing the right mouse button will change the orientation of the pointer. This pointing system does not rely on information about the 3D scene and is therefore easy to implement as an ‘add-on’ to current 3D systems, including live stereoscopic video.

The ‘Torch’ pointer is an image of a torch (flashlight to some) with a slightly diverging, yellow light beam which ends on the surface of the nearest object. Left-right and towards-away mouse movements are translated into rotation of the torch around the vertical and horizontal axes, respectively. The position of the torch in 3D space is altered by moving the mouse while pressing a mouse button; (left button: positions in a plane parallel to the screen surface, right button: positions in depth). This pointing system relies on detailed information about the surfaces of the objects in the picture.

Since two of the pointers need information about the surface of the objects their use is restricted to computer-generated images like *Virtual Reality* scenes. Therefore, all three systems were implemented on a fast graphics computer. A list of pointing requirements and hypotheses with regard to the three pointers was set up for testing:

The results of the tests are shown in Table 2. They testify that, generally speaking, the subjects positively evaluated all three 3D pointing systems. One obvious result is that for non-moving objects the ‘Floating pointer’ was evaluated as better than expected. The authors cannot recommend any one pointing system as such, but that the choice of pointing system is made according to the specific task or application. However, the authors affirm their main finding that it is possible to translate 2D pointing into 3D pointing for use with 3D PC applications and 3D video applications.

Table 2 Hypothesis with regard to pointing requirements

Pointing requirements	Surface pointer	Floating pointer	Torch
1. Accurate pointing	easy	difficult	easy
2. Hit objects in depth	easy	difficult	easy
3. Point at positions in space	impossible	difficult	impossible
4. Point in any direction	difficult	easy	easy
5. Point at transparent objects	easy	difficult	impossible
6. Point through transparent surfaces	impossible	difficult	easy
7. Move along any direction	difficult	difficult	impossible
8. Follow any moving object	easy	difficult	easy
9. Move along any object	easy	difficult	easy
10. Mark any object with a tag	easy	difficult	easy

4 Design for all – support for elderly and disabled

An area that is now receiving increasing attention is the accessibility to informatics and telecommunication systems (ITC) by elderly and disabled people. The Commission of the European Communities (CEC) has mandated a number of research projects and standardisation activities to rectify this situation. It is not acceptable in a modern society that up to one quarter of the population cannot use common telecommunication and information services.

With increasing standards of living and improving medical treatment, people live longer and are in better health than ever before. With the general lowering of retirement age coupled with larger pensions, people are better off economically and have longer active retirement time than was the case only a generation ago. These senior citizens make up the most important new customer groups (e.g. the fastest growing crowd of Internet users in the USA are retired senior citizens who have more time and money to spend than many other groups).

There are also large groups of people with various impairments, which are not as visible as the ‘classical’ disabilities such as deafness, blindness or motor impairment. These are people with learning disabilities, people with reading disabilities, people who are educationally disadvantaged, people with psychiatric diagnoses and people who are culturally or socially disadvantaged. These constitute a very large number of people, who have mostly gone unrecognised in industrialised countries and who have received little attention compared to people with the ‘classical’ disabilities, which often have powerful organisations and spokespersons.

The integration of disabled people into ‘mainstream’ society is gaining momentum in many countries. One fundamental requirement for independent living, however, is to have open access to basic information and communication services on par with all other citizens. Some people with profound impairments will obviously require special equipment and services, but very large groups of disabled people and senior citizens will only require marginal improvement of standard equipment and common services: e.g. larger keys, more legible markings and displays, improved sound quality and simpler operating procedures. It is a deep irony that despite, or maybe because of, the many recent technological advances, usability and perceived quality of most equipment and services have deteriorated in several respects; key-pads with reduced tactile feedback have become smaller and more 'fickle' to use; sound quality has worsened (with low quality synthetic speech increasingly replacing natural speech); displays are often added without any regard for legibility (with low contrast, small characters, obliterating reflections, inconvenient positioning, etc.); operating procedures are not very intuitive or much too complicated for many people; operating instructions are either absent or too difficult to read and follow for many people; the list can be carried on and on.

Recent legislation concerning people with impairments (e.g. *Americans with Disabilities Act* of 1990 and *Telecommunications Act*, now going through Congress) has put these issues in focus, especially in countries like USA that have the most liberal markets, and it is now only a matter of time before we will see similar legislation passed in Europe (either as national laws or issued as Directives by the Commission of the European Communities).

However, economic concerns may soon dictate as much as any legislation that something has to be done since increasing market shares of elderly and disabled customers will be significant and cannot be ignored. The tools already exist for justifying the positive cost/ benefit in improving the user interface, not only for ordinary customers but equally so for disabled and elderly customers [45].

4.1 Accessing the WWW by telephone

The Graphical User Interface (GUI) has been a catastrophic step backward for blind and severely visually impaired users of IT. Inherently, it is very difficult to convey all the information in a GUI screen in a meaningful way to a blind user, whereas the earlier line based systems were relatively easy to convert into Braille on an 80 character Braille reading device.

The World Wide Web (condescendingly referred to as the World Wide Wait) is emerging as the universal information source providing information normally found in books and periodicals, or by attending courses and conferences, etc. This is of great benefit to sighted people, who have instant access (well, nearly) to a vast amount of information. Blind people, however, are denied this source of information since most extant web browsers are GUI-based and thus inaccessible. Besides, it is necessary to have access to a computer for browsing the web, which is not common for blind people. Current Braille and speech screen readers are deemed quite unsatisfactory [46,47,48]. The idea of a speech-based interface for accessing the web via the telephone has been proposed several times, but little progress has been made up to now.

Zajicek & Powell [49] of the Oxford Brookes University, UK, have addressed this problem. The authors declare that the cognitive demands of browsing the web with a graphical user interface are considerable and require total attention. They envisage a future in which the requirements of continual update of information (e.g. when travelling, when shopping, or while performing leisure activities) increases the need for individuals to update themselves in situations where they have no access to a computer. They thus imagine many situations where it can be more useful to access information on the web via the telephone (e.g. a mobile phone) instead of being tied down in front of a computer.

To meet this challenge, the authors built a helper application they called *WebChat* that intercepts the HTML code that makes up a web page and provides a conceptually organised version of the web page using speech. The authors re-described the structure of a web page so that it could be 'read' out to a visually impaired user over the telephone. This involved forming a conceptual model of the information content of the page and its interrelationships and presenting it as speech.

A web page can be presented conceptually as a collection of textual objects and links. *WebChat* provides access to conceptually different parts of the web page using function keys and different synthesised voices to indicate conceptually different text features, such as headings, links and text paragraphs.

A graphical interface is essentially concurrent; i.e. many items of information are displayed simultaneously so that the user can engage in any one at any time. Speech based information is essentially sequential. The authors have devised a concurrency similar to that of a graphical interface by conceptually structuring the page to offer options such as lists of headings, lists of links and lists of keywords found in the text, or to read out paragraphs of text, etc. This they achieve by using a menu-bar, which is accessed by speech or function keys.

Zajicek and Powell have carried out experiments to evaluate the menu-bar as a tool for orientation on the web page and to determine which menu-bar option, headings, links or keywords, was perceived as most effective by the users. From their initial experiments the authors believe that there is a potential for the presentation of web pages in audio format. Although orientation and information retrieval took longer, the models created by the users from the available information seemed comparable with the models achieved by visual scanning of a web page. Users can become impatient if they have to listen to output they are not interested in. By providing the menu-bar for navigation the users can jump around on the web page. The authors are confident that if they can find a way of providing concurrency of information that is transparent to the user and embodies the power of the concurrency inherent in GUI, access to computers by telephone will open up an important new information channel.

4.2 Communication for non-speaking users

People who cannot speak are dependent on various forms of assistive technology such as hand-held and portable communication devices (PDAs) for communicating. Such small devices usually have low processing power and a link with a more powerful computer, e.g. for word prediction in text production, is desirable. A link to a powerful computer, e.g. over a telephone or mobile phone, would in effect give the user the processing power of a large computer in a cheap small hand-held device. What is the minimum bandwidth requirement for using a word-prediction application before a user is significantly delayed using a word prediction application?

Beattie, Hine & Arnott [50], the Micro-Centre, University of Dundee, Scotland,

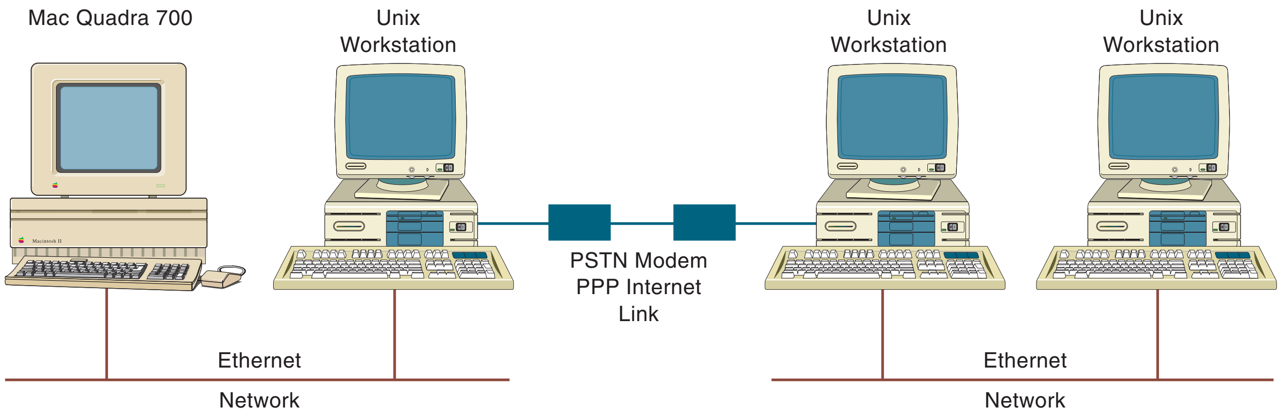


Figure 11 The basic experimental set-up used for all conditions in assessing the assistive process (after [50])

have performed a study to compare the influence of different baud rates of a telecommunication link between a local terminal and a remote computer on the usability of a text prediction application, and to compare these to the use of text prediction as a local application.

Rather predictably, they found that the interval between the key press to receiving the prediction rose as the speed of the connection fell (see Figure 11). The delay at the two slowest baud rates (1248 ms at 1200 baud and 800 ms at 2400 baud) of the remote applications were unacceptably long when entering text and waiting for a prediction to be displayed. The authors conclude that word prediction over a remote link can be used effectively if the transmission rate is 9600 baud or higher. Below this rate the user notices the delay which leads to frustration and errors. The large processing power required by some assistive techniques may thus be available to users by a cheap, small portable device linked to a remote server, provided the link has sufficient speed to satisfy the use of the application.

4.3 Picture annotation for non-speaking people

As telecommunications have become more and more important, people with speech impairments are even more disadvantaged than before. In a face-to-face situation, a non-speaking person can always resort to writing, pointing, pantomime or pictures. In telecommunications, a non-speaking person has a much more restricted variety of communicative techniques. The conventional telephone provides a multitude of interactive services (e.g. 24-hour banking, and other services) which do not require speech input. Many non-speaking people regularly use text telephones and fax services, but studio or desktop video-telephones are also being used as an alternative to face-to-face meetings. The Internet and multimedia services hold high promise for non-speaking people.

The basic problem, however, is that text input instead of speech is very slow. In conversation and communication pictures and graphics can be used to convey complex ideas and various symbol and picture systems have been devised for Augmentative and Alternative Communication (AAC).

To test the practicality of graphic communication as an alternative to speech, Hine, Beattie & Arnott [51] of the MicroCentre, University of Dundee, Scotland, compared the time taken to exchange information in text form and in picture form. Their hypothesis was that some users could exchange information at a significantly faster rate by displaying and annotating pictures than by typing the same information.

The experiment involved a scenario where eight subjects indicated to a remote ‘Travel Agent’ a number of places they wanted to visit in their holidays, using either a text telephone (simulated on a computer) or by displaying and annotating a map on the computer screen.

In a preamble, the ‘Travel Agent’ first asked which state was to be visited. The reply was either a name typed on the ‘text telephone’ or the display of a state map on the terminal screen. This was followed by a request for the first town to be visited. The reply was either a typed name or the positioning of a box marker on the map to indicate the desired town. The ‘Travel Agent’ requested confirmation by repeating the chosen name in a text window, and the subject confirmed by either typing “yes” or placing an arrow in the box on the map. If the ‘Travel Agent’ responded with an incorrect name, the subject either retyped the correct name or again placed the box marker around the desired town on the map.

Five experimental episodes where the ‘Travel Agent’ asked for the next town to be visited and the road to get there followed this preamble. In each case the subject either responded by typing the names via the ‘text telephone’ or by indicating on the screen map. The time to complete each episode was recorded. The full experiment consisted of eight subjects doing four sessions each with two text and two map conversations. All subjects had the same eight scenarios and the text and map sessions were counterbalanced to avoid any order effects.

The results are very variable and can only hint at any tendency. Five of the eight subjects completed the text tasks faster on average than the map tasks. However, the eight subjects varied with respect to speech and upper limb motor impairments, and used various screen pointing devices (mouse, roller-ball, chin controlled roller-ball, head-switch driven mouse and keyboard, on screen keyboard, concept keyboard and *Keyguard*) making any comparisons and conclusions invalid. Although the authors explain some of the variability and find consolation in the fact that at least three subjects did better with pictures than with text, supporting their original hypothesis that "some users can exchange information …", the results must actually be regarded as highly inconclusive due to flawed method and bad experimental design. It should be interesting to see results for text vs. picture communication that are truly comparable, with the right screen design it may well be that communicating by pictures rather than by text is a better proposition for many users and applications.

4.4 Disabilities and telecommunication services

The use of a telecommunication service involves interaction with a complex set of elements and performing a variety of activities. Modern and future telecommunication services will allow using a variety of media, both in the control of the services and for the information exchange. If the user has a disability the service may be unusable. The barrier may lie at three levels; in the *terminal,* in the *network* or in the *service* (see [52]).

Hine, Beattie & Arnott [53], the MicroCentre, University of Dundee, Scotland, present an interim attempt to develop a method of anticipating and recording the problems that users with disabilities face when using telecommunication services and the potential solutions that could address their problems. Their analysis is part of a wider issue of finding services that users need and the need for adapting current and future services to the users' needs. Their analyses and procedures for assessing user requirements and proposing usability improvements of present and future telecommunication services derive from the ACTS UMPTIDUMPTI project and follow accepted mainstream methodology for usability assessment. Similar studies and analyses have been carried out both in past and present projects (e.g. TUDOR, COST 219, TIDE, Interact, Include, ETSI PT6V, etc.) without much impact on service operators, manufacturers or regulatory authorities; new services and equipment are still being introduced with no regard for people with impairments. The quality of work in a study is no guarantee for its impact on the practices of the operators and the markets, the findings and results are duly acknowledged, but nothing happens to alleviate the situation for the disabled user.

The authors conclude that the process outlined in their report allows user requirements for the usability of telecommunications to be analysed, problems to be highlighted and solutions to be suggested and investigated so that issues can be flagged for a variety of users under consideration. In general, this method illustrates the wide range of factors that may influence the usability of a telecommunications service, particularly for users with disabilities.

4.5 Mobile access to WWW services

A GSM telephone connected via a PCMCIA card to a personal (laptop) computer moving about in a cellular radio network is a technology that will enable 'mobile' access to the Internet. This may provide a number of novel opportunities, such as library consultation, distance education and banking services, to people who are mobility or motor impaired. These benefits, however, can only be realised if the equipment is fully accessible to the intended users, i.e. any user barriers in the terminal, network and service must be removed as early as possible in the design stage.

Van Mele, Edwards, Verheust & Spaepen [54], Katholieke Universiteit, Leuven, Belgium, wished to determine the feasibility of mobile access to the WWW by mobility impaired people using a laptop computer connected to a GSM handset. The authors believe that mobile telephones will serve an important role as assistive technology for mobility impaired people, i.e. people who are bound to a wheelchair, who have difficulties leaving their home or who cannot easily access a library or come to classes.

The authors used a system consisting of a Nokia 2110 I GSM mobile telephone and PCMCIA card connected to a T400 CDT notebook laptop computer. Ten students (aged 19 to 32) with various motor impairments served as subjects. All subjects had previous experience with computers. Where there was GSM coverage, the tests were held in the students' rooms, otherwise a suitable location was found. In the first test the subjects were asked to set up the mobile equipment, write an e-mail off-line and send it using Pegasus e-mail application. In the second test the subjects had to search for some specific information on the Web using Netscape® navigator. The time to complete the various tasks was logged by the computer and the tests were video filmed with time registration.

The time needed to connect the GSM handset and install the PCMCIA card in the laptop varied from 40–60 seconds for subjects with no upper limb impairments to 80–100 seconds for subjects with more severe upper limb impairments. Most subjects reported that a built-in PCMCIA card would significantly simplify set-up of the system and make the terminal more accessible.

The sending of the e-mail showed that the quality of the mobile network connection (which varied from 0.5 to 4.0, with a mean of 2.5) was a deciding factor. The second test required an even higher network quality. All subjects complained about the noise due to the interference between the GSM handset and the speaker output of the notebook during data communication. The audio output and PCMCIA slots were only 10 cm apart and the GSM handset cable was too short to have the GSM phone at sufficient distance from the audio output. More serious, though, was that normal control of the cursor was lost if the mouse or mouse emulator was on the same side as the GSM handset, especially during long connections with the Web server. The handset had to be moved to the opposite side and at least 40 cm away from the laptop before normal control of the cursor could be resumed. Also, interference with the GSM handset affected the operation of a hearing aid and arm prosthesis.

The task of writing and sending an e-mail using Pegasus was completed in between 181 and 685 seconds. The task of searching for information on the Internet using Netscape was completed in between 122 and 673 seconds, while the number of selections necessary to find the required information varied from 13 to 100. The high values are mainly attributed to one subject who in addition suffered from tunnel vision and thus had great difficulty reading the screen. Post-test questions revealed that 90 % of the subjects favoured using these applications professionally.

The authors conclude that the trials have demonstrated the feasibility of mobile GSM/laptop set-ups for people with motor disabilities. However, the tests have also revealed significant user barriers on all three levels; terminal, network and service. It will be necessary to reduce interference and simplify operation of terminal equipment, to improve GSM coverage and to speed up access to services to make mobile access practical for motor impaired users.

4.6 Video quality for sign language communication

Video communication offers great possibilities for deaf people using sign language. For selecting or developing video systems for communicating in sign language, a method is needed to determine the practicality of any system.

A simple method of measurement was developed in the project "Quality requirements on video communications for sign language" and is reported by Hellström [55] of Omnitor, Sweden. The method extracts technical factors related to the requirements of sign language users. The method is easy to use and requires only one or two video tape recorders (VCR) and a test tape that was produced within the project. Although the method is based on sign language, it does not require knowledge of sign language to be used. In use, the videotape is played over the video telephone system to be assessed (see Figure 12). It is preferable to record the outcome at the far end, but it is possible to make an assessment by direct observation.

The videotape contains nine video sequences designated:

1. *Silent voice* (reference video with American Sign Language)
2. *Fast* (rapidly signed language)
3. *Slow* (slowly signed language)
4. *Frame rate* (picture update rate, >12 frames/s, 20 f/s is preferable)
5. *Static resolution* (striped still patterns for resolution measurements, QCIF (176x144) acceptable, CIF (352x288) good)
6. *Blur* (sequence for judging blur of details in sign language)
7. *Sync* (sequence for measuring synchronisation of sound and video, should not be higher than 0.1 s for lip-reading supported by audio)
8. *Lip reading* (sequence of talking person to determine usefulness for lip-reading)
9. *Delay* (sequence for measuring picture delay (sound to be sent trough separate cable) delay should be less than 0.5 s for smooth conversation and below 1.2 s to be acceptable).

The validity of these quality factors has been verified in conversation tests by proficient signing persons. Using this method, it should be possible to ascertain the minimum quality necessary in a prospective video telephone system for sign language communication.

In a comparison of projects using different techniques, bandwidths and services, Delvert [56] of Kalejdo Management Consultancy, Sweden, investigated the merits of different systems for signing deaf people, including comparison of text telephones and video telephones at different frame rates and the usefulness of text telephony relay services. Although the results reported are somewhat inconsistent and fragmentary, a general picture emerges showing that deaf signing people may benefit from all the systems tested, and that the usefulness of any video system improves with increasing bandwidth, i.e. higher picture frame rates and/or higher spatial resolution. The issue of deaf people, either pre-lingual deaf sign language users, or adventitiously deaf people who can speak, read and write, using telecommunications has been subjected to many research projects, which all tend to gravitate towards the same results as presented above.

5 Group working support services

Group working has emerged in recent years as a novel way of working together on a document or other task via the telecommunication network. The actors taking part in group-work activities may be geographically proximal or dispersed, this is not the main issue – the main issue is that the system and the services must support interactivity between several people dedicated to solve a common task via the network. Special forms of group work in distributed organisations are telecommuting, mobile working, workgroup computing, videoconferencing, etc. In a companion article in this issue of *Telektronikk* (see *The ergonomics of teleworking*), Bakke deals more in depth with these matters.

5.1 Assessing user needs for future work practices

To assess the user needs of mobile professional workers, Sallnäs, Hedman & Virtanen [57] of Telia Research, Sweden, studied selected members of mobile pro-

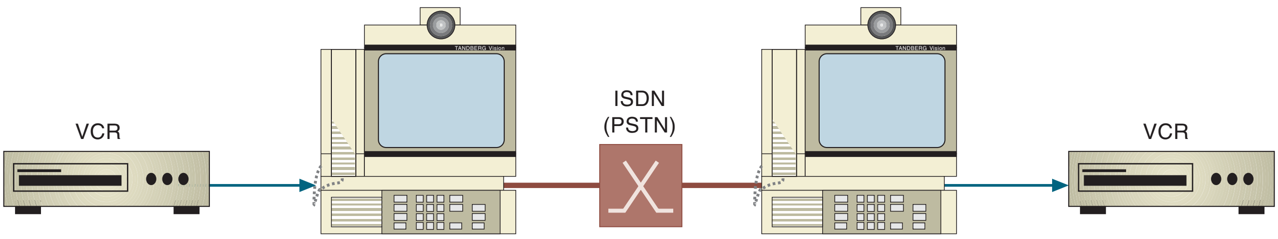


Figure 12 The basic experimental set-up used for the quality measurements (after [55])

fessional teams in demanding environments; a fire fighter, a building worker, a plumber, a salesman, and a reporter and a production manager from a local TV station. The topics studied were social group processes, awareness of other members of the organisation, instant team status, applicability of group software, etc.

The results of these studies show that multiple methods are necessary to obtain an understanding of present and anticipated work practices. The studies also provided information about the most important patterns for communication, co-ordination and co-operation. The results from mobile workers showed that this user group mostly worked synchronously, either face-to-face within teams or via various audio communication systems that support awareness of colleagues' geographical position, status of work processes, etc., but working distributed from other functions in the organisation. For all user groups, the present mobile communication was regarded to be the most important working tool, since the communication system is the only tool that supports real-time access to other functions in the organisation.

Telecommuters seemed to fall into two distinct categories: those who planned their telecommuting work and those who see it as an opportunity to work *ad hoc*. The first group defined special telecommuting days and planned specific work tasks for these days. They regarded their office days as social, devoted to meeting people and attending meetings. The *ad hoc* group, on the other hand, avoided planning and tele-worked when the need arose to execute individual tasks. They also often split days between office work and tele-work. They also tended to work more in the evenings and weekends. Both groups organised their work at home and at the office in very similar ways – most of the working time was devoted to computer work and asynchronous communication, i.e. e-mail and fax.

An unexpected finding was that, even though management tended to be positive to telecommuting, the colleagues of telecommuters were often critical, fearing isolation and complications in their own work as a result of introducing telecommuting in the organisation.

The authors conclude that the tests had revealed two important categories of functionality in the five professional groups studied (see Figure 13):

- All groups utilise information, both from internal and external sources. Today, information is mainly communicated to mobile teams via auditory channels, but there was a desire for more visually displayed information. Generally, these mobile professionals had only little need for processing information.
- There was a general desire for awareness through shared contexts (Instant Team Status). This is now mainly achieved through auditory channels, but many professionals also had visual co-ordination aids that support awareness of team status, such as work process charts, geographical distribution charts and maps. Team members desired resources, which could be used instantly to solve time constrained tasks, especially visual and auditory mobile access to information to verify the team status.

The authors also conclude that informal communication in distributed teams is not well supported by extant technologies. Identifying and dealing with such factors is necessary for the successful implementation of future multimedia services for mobile professionals.

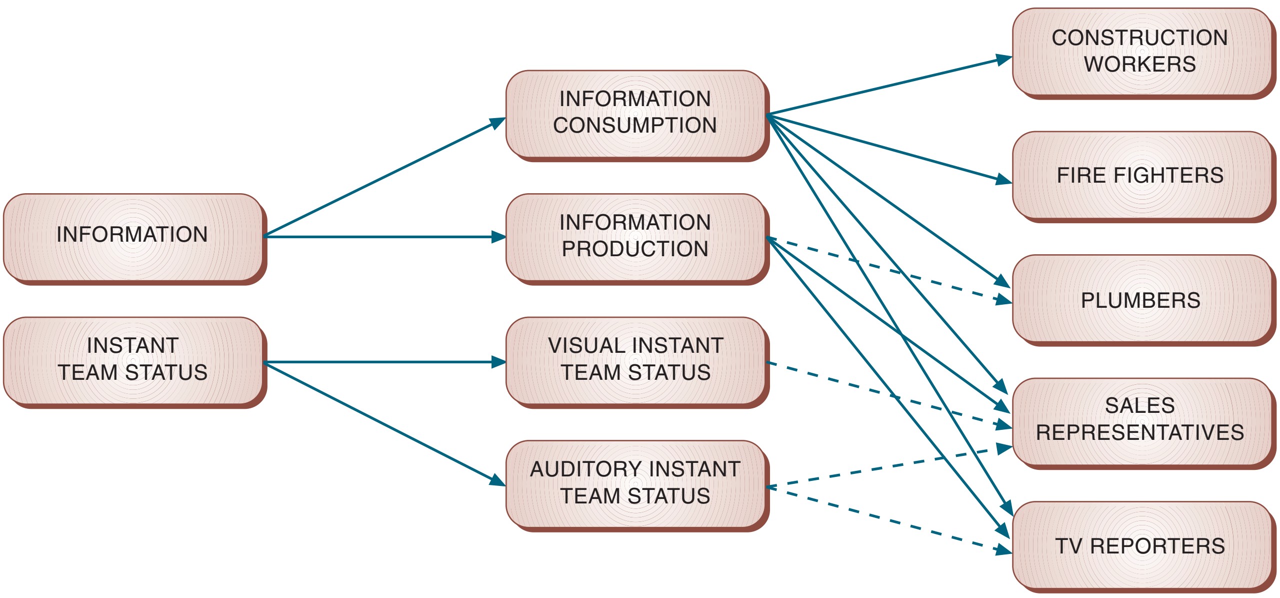


Figure 13 The diagram shows the relationships between User-groups *and* Functionality *in their current work-situation (after [57])*

5.2 Collaborative tools in distributed research projects

International research projects require a high degree of communication and collaboration over long distances, and work often has to be done to tight time schedules. Efficient communication systems have thus become a prerequisite in such work. In what way synchronous and asynchronous multimedia communication tools can support work in distributed research teams was studied by Böhm*, Corte#, Marion#, Oberndorfer*, Schmitz* & Uellner* [58] from *Deutsche Telekom, Technology Centre, Darmstadt, Germany, and #Telecom Italia, CSELT, Torino, Italy. They find that even e-mail (if available) seems to be insufficient for the exchange of complex information, and face-to-face meetings are too costly in terms of travel time and expenses. However, multimedia collaboration tools, if properly chosen, seem to be a solution to these problems.

Since project members play different roles in a project, such as project participant, task leader and project leader, they also perform different tasks, and a number of different activity classes were identified, such as work process, decision making, design, discussion and co-ordination. Figure 14 gives an overview of the interrelations between roles and tools. The authors wished to determine which capabilities of collaboration tools could support which roles and tasks. They also wished to know the exact requirements for communication in a distributed research project.

The authors dealt with this by distributing two questionnaires among the members of R&D departments in the EURESCOM *(European Institute for Research and Strategic Studies in Telecommunications)* community. The aim of the first questionnaire was to obtain information on:

- What kind of telecommunication activities do users carry out in the context of their EURESCOM work
- What kind of computer support is available to users today
- What kind of computer support do users think would help them with their daily communication problems.

The second questionnaire had three parts asking information about:

1. Personal data and characteristics of the EURESCOM projects people were involved in
2. Current and expected future use of collaboration tools, classified into the following 10 groups:
   1. *Decision Support*
   2. *Desktop Conferencing*
   3. *Document Management*
   4. *File exchange*
   5. *Joint Editing*
   6. *Joint Viewing*
   7. *Message Conferencing*
   8. *Scheduler*
   9. *Shared Blackboard, and*
   10. *Workflow Management*
3. Inquiry as to whether specific combinations of tools addressed under Part 2 have synergistic effects, and expected use of tools for specific classes of collaborative work.

One important finding was that nearly all respondents used the telephone, e-mail and face-to-face meetings almost daily. Fax, file transfer, postal mail, exchange of diskettes, Internet and bulletin board was used weekly or less frequently, while video telephony, video conferencing, videotext, voice mail and computer-supported telephony was hardly ever used.

The authors' main findings were that over three times the current number of users expect to be using *Joint Viewing* and *Joint Editing* tools in the future, although expected frequency of use was not expected to change significantly. Twice the current number of users expects to use *Shared Blackboard, Workflow Management, Decision Support* and *Desktop Conferencing tools*. Respondents estimated that they will use *Shared Blackboard* and *Desktop Conferencing* tools more often in the future, while *Workflow Management* and *Decision Support* tools are expected to be used less frequently. The number of users of *Message Conferencing*, *File exchange*, *Document Management* systems and *Scheduler* tools is not expected to change significantly in the future, and the frequency of use of *Scheduler* tools is even expected to fall (see Figure 15).

In conclusion, the authors remark that members of distributed research and development teams require and are willing to use collaborative tools in order to improve the effectiveness of their joint work efforts. The tool classes that will be of greatest importance are *Desktop Conferencing, Joint Editing, Joint Viewing, Shared Blackboard* and *Workflow Management*. The authors find the respondents' modest confidence in *Document Management* tools a bit surprising. Obviously, many users are satisfied with their e-mail and *File Transfer* tools and they do not see any urgent need for alternative systems.

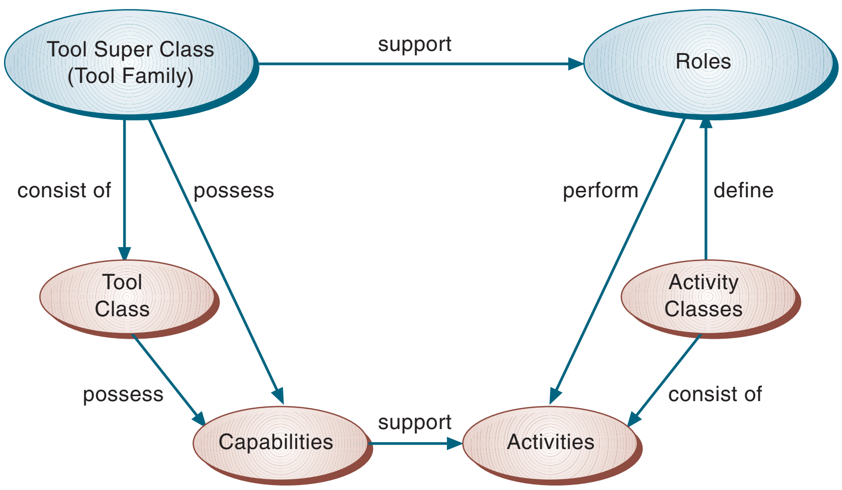


Figure 14 The diagram shows the relationships between Roles *and* Tools *in the distributed research project (after [58])*

5.3 Virtual office environment with shared room awareness

One advantage of working from a home office is to be able to work in seclusion and not be constantly observed by colleagues, while this same isolation when working at home, away from colleagues, is often quoted as the main drawback of working at home. How can the tele-commuter have the best of both worlds, i.e. both to revel in total seclusion when that is most appropriate and to enjoy the social awareness of colleagues when that is more desirable?

To be able to strike a balance between these two incompatible wishes, Honda, Tomioka, Kimura, Ohsawa, Okada & Matsushita [59] of Keio University, Yokohama, Japan, have realised a system which enables home workers to have a home office in 3D virtual space and work in it together with their virtual colleagues. The system provides a virtual, shared room on a Sun station and provides simulated awareness of other colleagues to aid communication. In a real office, we are able to sense other colleagues' awareness of us, such as actions and gaze, and other occurrences in real time while working.

To realise awareness, the home office was provided with two interfaces called *Around View* and *Sound Effect*. In the real world we are able to sense if someone next to us sits down or stands up, even though we do not look directly at the person. This is because our field of vision extends more than 180°, while the field of vision where we see details extends only about 60°. The *Around View* interface is based on the human visual field, i.e. the field of vision subtends 180°, but only the middle 60° depicts any detail while the outer flanks gradually show less and less detail. With *Around View* users can therefore have awareness of a person's motion in the side fields but detailed picture of a person sitting in front of them. To see someone at the side in any detail, the user has to turn, but this information is shared by the other office workers, who will be aware that they are now being seen. A sensor on the home worker's swivel chair sensed the rotating angle, and if this exceeded 30° the view of the office also rotated to the same side.

In the real world we also have acoustic awareness information. This consists both of human utterances, both voluntary such as speech, and involuntary, such as sneezing, and environmental sounds such as the movement of people around you, chairs being moved, office equipment such as printers and copiers, and background noise such as ventilation systems, traffic, etc. With *Sound Effect* added, users working at home should get a more realistic awareness that other people actually work in the office.

The study consisted of a number of tests, involving many different methods, including video filming and recording of EEG (electro-encephalography) brain waves to determine degree of intellectual concentration from viewing the video and analysing the EEG for presence of beta-waves. The size of the virtual awareness space could be varied in three levels, Narrow, Normal and Wide. The level was determined by and varied inversely with degree of concentration, the more the worker concentrated on the task at hand, the smaller the awareness space (see Figure 16).

The authors conclude with an informal evaluation of the system on the following five points:

1. Presence of the user
2. Presence of the other workers
3. Facility of communication
4. Position awareness
5. User interface.

The authors show that the system provides the expected awareness level for the task at hand. When little 'social contact' is required while working on a difficult task the system automatically reduces awareness level to *Narrow*, but it returns to *Normal* awareness in an 'ordinary' work situation and changes to *Wide* awareness when it senses that more 'social contact' is preferred.

5.4 Tele-working in Rome: a case study

People in the city of Rome spend an average of three weeks or more a year commuting to work. Most of these

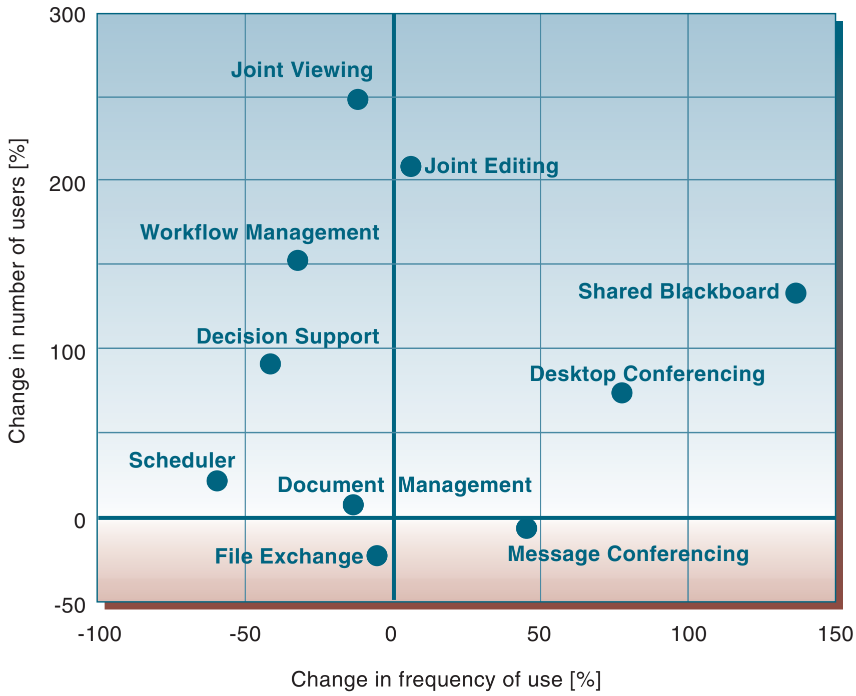


Figure 15 Expected changes of usage. The expected percent change in the number of users (ordinate) is plotted against expected percent change in frequency of use (abscissa) for ten common collaborative tools applications (after [58])

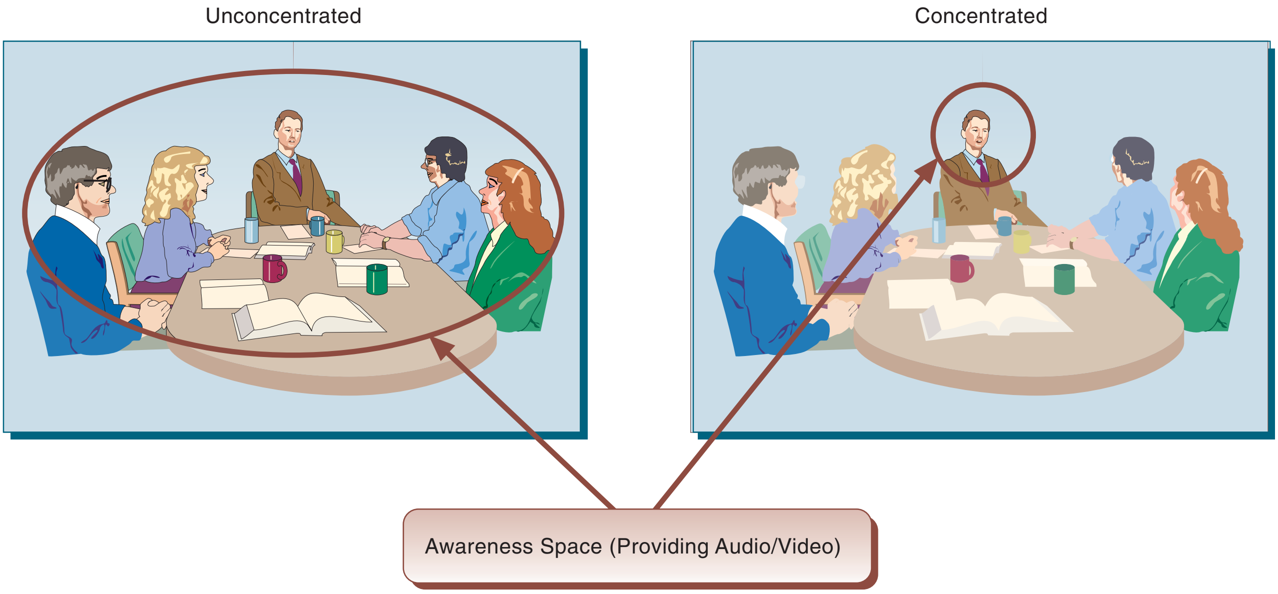


Figure 16 The experience of shared-room awareness space changes according to the tele-worker's mindset: When the concentration on a task increases the perceived awareness space shrinks to the proximal personal area (right panel), but if the mental concentration diminishes the perceived awareness space expands to encompass the whole room (left panel) (after [59])

people are 'information' workers. This begs the classical question: "why do these people have to go to an office 30 or more kilometres away to work on a personal computer or talk to a telephone that could be in their homes or in an office much closer to their homes?" [60,61].

In Rome, with nearly three million inhabitants, some 5,800,000 commuter trips, average length of 10 km, are made every working day. Preliminary results from field trials with tele-working in the municipality of Rome, the *Rome Traffic Decongestion Tele-working Project* (Rome TRADE), is reported by Matarazzo [62] from Fondazione Ugo Bordoni, Rome, Italy. The main objectives of the Rome TRADE project were: to identify characteristic commuter origin/destination patterns; to carry out tele-work trials among employees of the Rome Municipality; and to develop a scenario for introducing tele-work in other public institutions.

Although the basic rationale for the project was to find ways to decrease traffic congestion in Rome, the author initially stressed the importance of planning tele-work from a human factors point of view. He finds that it is just as important to focus on the suitability of products as it is on process, performance criteria, productivity or establishing trust. On the basis of initial 'scouting', 35 employees from ten different municipal units in the centre of Rome were selected to participate in a project lasting three months. A flexible tele-work scheme allowed the participants to work either from their homes or from a local tele-working centre. Data was collected via questionnaires administered after each tele-working day and after two or more traditional office days.

Based on a sample of only 18 participants, some preliminary results are reported. The average frequency of tele-working was about four days per month. The number and length of commuting trips did decrease with tele-working, from 3.7 trips/non-tele-working day of average 30 km each, to 1.6 trips/tele-working day of average 25 km each. The results show no significant effect of gender, age or profession on the frequency of tele-working. However, a most striking effect of the *direction* of commuting on the frequency of tele-working was found: People who commuted from the centre to the periphery, or outskirts, of Rome tele-worked significantly more often (average 7.3 days/month) than other groups, while people who commuted similar distances from the outskirts to the centre chose to tele-work less often (average of 4.1 days/month).

The overall number of interpersonal communications decreased with tele-working from an average of 15.9 to an average of 12.7. The number of communications with superiors and collaborators remained virtually unchanged in tele-working. While the number of communications with colleagues and clients decreased, the number of communications with friends and relatives increased.

All tele-workers rated the tele-working situation as more *interesting,* more *pleasant,* more *useful* and more *productive* than the office situation. The technology caused few problems: Most tele-workers communicated by telephone, fax or computer e-mail, and used a computer for most of their work either at home or at the work-centre.

Autonomy in performing the job tasks increases significantly with tele-working, the tele-workers become more independent and ask less for advice, so even though tele-workers interact less with superiors and collaborators, their decision making is not impeded.

Finally, Matarazzo reports that the feeling of isolation in tele-workers was *not* significantly changed in the tele-working situation. This runs counter to conventional wisdom that isolation, both socially from colleagues or as lack of information about what is going on in the organisation, is one of the most often proffered reasons against tele-working. The finding may partially be explained by the relatively low frequency of tele-working in this study, but the preliminary results from the Rome TRADE project do dispute widely held beliefs among opponents of tele-working.

5.5 Broad-band technologies in tele-education

Tele-education, which is a special form of distance education, shares many features with tele-working: i.e. geographically distributed people working from their homes interacting with tutors via telecommunication equipment and services. The main perceived barriers to a broader acceptance of tele-learning and tele-teaching are related to various problems in interacting with technologies and to general resistance in teachers and students alike, especially in the fields of the humanities and the social sciences. Different virtual classroom situations may be created to suit different topics and the different needs of pupils, from the simple use of e-mail to the use of very sophisticated multimedia services using speech, music text, drawings, animation and video to create so-called 'tele-presence' to support contact between teachers and pupils.

To investigate the effect of various multimedia systems on the learning climate in tele-education, Papa & Spedaletti [63] from Fondazione Ugo Bordini, Rome, Italy, carried out experiments using a broadband network to emulate virtual classroom situations in the laboratory.

In configuring the virtual classroom multimedia system, some important features were aimed at:

- Possibility of self-view
- Continuous audio presence
- Continuous video feedback to the tutor
- Suitable video signal switching procedures
- A 'multimedia chair' for the tutor.

Nine students (aged 21–25) attending a course in Communication Mass Sociology from the University of Rome "La Sapienza" served as subjects. The course consisted of six lectures given by tutors using the 'multimedia chair'. The students were distributed in three rooms that communicated with the tutors' room. After each lecture, a questionnaire was issued and the subjects were asked to rate various aspects of the learning climate, various usability factors and finally, to give an overall evaluation.

The results reported by Papa & Spedaletti show that the subjective ratings of the learning climate, although they were generally high, did not change significantly from the first to the last lecture. The usability ratings were also generally high, and did not improve significantly over the six lectures. The same was also the case with the overall ratings by the students, no significant improvement, but ratings were generally high.

The teachers judged positively the procedures centred on their needs, and the teachers designed their lectures using most of the opportunities offered by the multimedia systems. The availability of different multimedia does change the tutors' use of the media. There were stronger time constraints in a virtual classroom than in ordinary face-to-face tuition. A reduced sense of presence among students compared to face-to-face situations was perceived by the tutors, and a virtual classroom demands developing new skills by tutors, e.g. being able to deliver a lecture in front of a camera.

In conclusion, the results indicate that no real differences were found in the students' evaluation of the usability dimensions over the duration of the experiment, although the system was generally positively evaluated. In general, the interactions permitted by the system were satisfactory. This kind of experiments are much too short and limited in scope to give any real evaluation of tele-education, classroom presence and suitability of multimedia systems in tuition, but they can pinpoint specific problems or features that can be used to improve the design of future trials.

6 Product development

The development process of telecommunication products and services has changed substantially during the latter years due to the application of various design strategies and the gradual introduction of various human factors procedures. These include such evaluating techniques as user participation, paper prototyping, software prototyping, heuristics, 'Wizard of Oz' simulation, cognitive walkthroughs, to mention only a few current methods. One user interface that has gained wide acceptance is the Graphical User Interface (GUI) combined with direct object manipulation, i.e. clicking and dragging icons to perform various actions that were previously achieved by text input and short commands. A well-designed GUI may substantially enhance the usability of ITC products or services, but just applying GUI indiscriminately does not necessarily guarantee a user-friendly product. It is still necessary to work through the various design process stages in a systematic manner and perform the required analyses and, not least, act on the results obtained to achieve a better product.

6.1 Clever GUI design gives value for the customer

The use of the Graphical User Interface (GUI) has spread rapidly from the world of the high performance computer to a host of products and services in the field of telecommunications, and several other areas. The same way a good GUI can enhance the usability of a product, a poorly designed GUI may be no better or even worse than the older method it replaces.

Fleet, Talbot & Winlow [64] from Nortel's Corporate Design Group, Ottawa, Canada, have developed a six-phase process for the design and evaluation of product interfaces with the clear objective of offering customers supreme value. Each phase in the process is designed to address a critical interval in the product development cycle and contribute to design excellence (see Figure 17).

The first phase is a *Domain analysis*, or to understand the problem space, i.e. the who, what, where, when and why of the product domain. To clarify such questions, various techniques such as site visits, interviews, direct observation, surveys and questionnaires can be used to identify significant aspects such as:

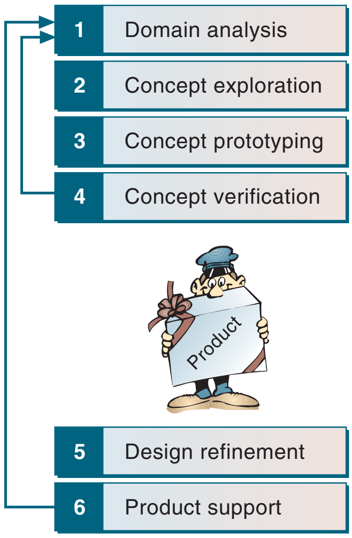


Figure 17 The six main phases of the design process. The first four phases may be iterated as often as necessary (after [64])

- current and future market trends and needs that may influence the interface
- current 'best-of-class' interfaces and their perceived customer value
- distribution channel variables that may influence the interface
- end-user profiles, tasks and requirements
- system requirements, features and development schedule
- current and future user interface technologies and techniques.

The *Concept exploration* phase is meant to generate ideas that may lead to fresh user interface concepts. In this phase it is essential to encourage unorthodox thinking. Basic techniques in this phase may include:

- Brainstorming to generate fresh concepts
- Organisation and grouping of concepts
- Voting between the best concepts.

In the third phase, *Concept prototyping*, concrete examples are used to demonstrate the value of the concept to the stakeholders. Prototypes can offer developers detailed feedback in a timely, cost-effective manner. The authors describe three levels of rapid prototyping:

- *Paper prototype;* using pencil sketches, bubble diagrams or flowcharts to portray tasks, object models and task flows
- *Walk-through prototype;* using animation, HyperCard™ stacks or QuickTime™ films to demonstrate functionality
- *Software prototype,* fully or partially functional prototype used for usability testing, demonstrations and design transfer.

The purpose of the next phase, *Concept verification*, is to minimise costs in terms of money and time by verifying concepts at an early stage rather than at later stages in the development. Usually, a number of iterations are required, and these are more cost-effective to perform as early as possible. Some common testing techniques are:

- *Focus groups*, good for collecting qualitative feedback from users
- *One-on-one interviews*, useful for looking at concepts in more depth
- *Usability testing*, effective with sufficiently developed concepts
- *Co-discovery learning*, a preferred technique for testing interactive products.

By the fifth phase, *Design refinement*, the design is essentially complete. Further usability testing for fine-tuning the concept may still be done and missing details can be filled in, but at this stage it is relatively expensive and technically difficult to make any significant technical changes.

The final *Product support* phase is meant to ensure that maximum benefit is derived from field trials and early product feedback. It is important to verify if the original design hypotheses were correct, which specific design features were most successful, and to seek recommendations for improving future products.

In conclusion, the authors present an example of how the six-phase design process was successfully employed to design a network managing system. Their studies have shown that over 50 % of product defects can be traced to inadequate understanding of customer requirements. Thus, the real measure of success of any design approach is that it substantially reduces the risks to the manufacturer while increasing the value to the customer.

6.2 Individual differences in design of automated services

How can individual differences in users' performance best be assayed and incorporated in the development of new services? This question was addressed by Love, Foster & Jack [65], from Centre of Communication Interface Research, University of Edinburgh, Scotland. They employed a three-step model proposed by Egan & Gomez [66]. Step 1 'assays' the effect of individual differences such as age, personality or cognitive skills on task performance. Step 2 'isolates' where salient differences identified in Step 1 affect interaction most. In Step 3 features which have been isolated in Step 2 are 'accommodated'. Using Steps 1 and 2, the authors performed an experiment to investigate the hypothesis that individual differences in subjects' performance can be attributed to individual characteristics, and that these differences can be isolated to specific sub-tasks in the user-interaction.

To test the hypothesis an automated music catalogue telephone service with menu selection was designed. The catalogue was hierarchically organised with menus of three to five items. The menus were arranged like the grouping of CDs in a typical music store: The first level *category menu* offered Blues, Jazz, Classical, Rock & Pop, and Folk, the next *artist menu* level offered a choice of four artists or composers for each category, the next *album menu* level offered a choice of three albums from the chosen artist/composer, and the last *track menu* level offered a selection of three tracks from the chosen album. 'Audio-on-demand' was provided at the *track menu* level, allowing the subjects to listen for up to 30 seconds to any track from the selected album.

Thirty subjects balanced for gender participated. The subjects interacted with the on-line music catalogue via a telephone keypad. The task was to compile a 'personalised' CD by ordering four tracks, two of which were specified in the instructions and two to be selected by the subject. Each subject attended four sessions, in which a new CD was compiled and one each of four psychometric tests was completed:

1. *NEO – Personality Inventory Revised* [67], measures neuroticism, extraversion, openness, agreeableness and conscientiousness
2. *AH4 Test of General Intelligence* [68], assesses general verbal, numerical and spatial reasoning skills
3. *The Paced Auditory Serial Addition Task* [69], measures information processing capacity
4. *The Rey Auditory Verbal Learning Test* [70], assesses both short- and long-term memory.

Two objective performance measures were also obtained – the *interrupt rate* and the *silence rate*. Interrupt rate is the ratio of the number of key-presses *before* the end of a system prompt to the *overall* number of key-presses made by a subject to menu prompts during an experimental trial. A 'silence' was registered when a response was not made before a 'time-out' period of 5 seconds after a system prompt. The silence rate is the ratio of number of 'silences' to the overall number of responses made during an experimental trial.

No significant differences in *silence rate* were found for any individual characteristics, including age and gender. Also, no significant differences in *interrupt rate* were found for personality, memory or processing ability factors. However, significant ($p < 0.001$) effects of age and verbal ability on the interrupt rate were found (see Figures 18 and 19).

The results suggest that differences in user behaviour with an automated service may be explained in terms of individual differences. Also, the experiment has shown that once the salient individual differences have been identified, it is possible to determine where these individual traits produce their greatest effect.

Since it was found that age and verbal ability contributed significantly to the variation, the next important issue to be addressed would be how to 'accommodate' these differences (Step 3 [66]). One approach would be to use speech input rather than key-presses. According to the authors, several of the older subjects actually expressed that they would have preferred to speak to the system.

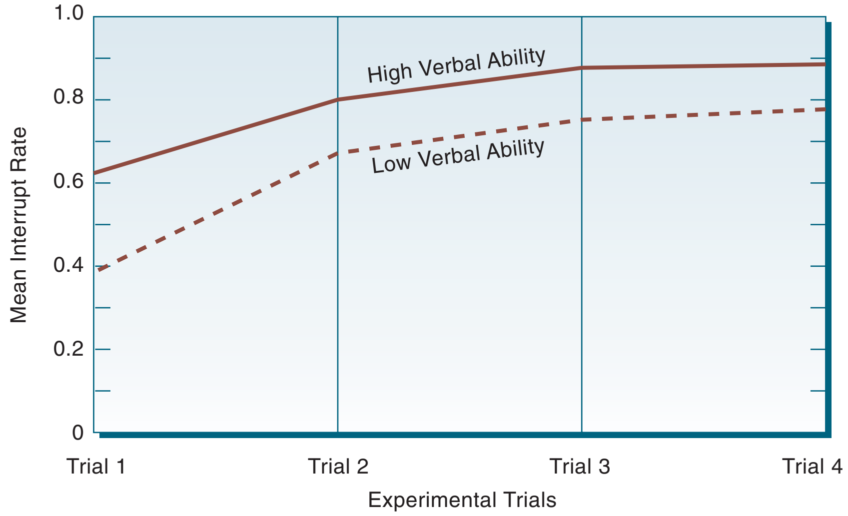


Figure 18 Mean Interrupt Rate *(ordinate) is plotted as function of trial order (abscissa) for two different verbal ability groups:* High Verbal Ability *(continuous line) and* Low Verbal Ability *(broken line). The differences between the groups were significant ($p < 0.001$) (after [65])*

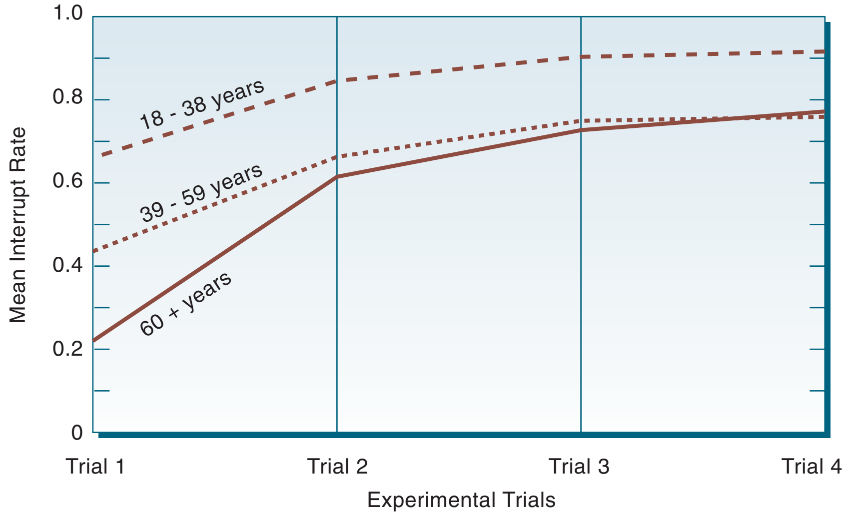


Figure 19 Mean Interrupt Rate *(ordinate) is plotted as function of trial order (abscissa) for three age groups: 18–38 years (broken line), 39–59 years (dotted line) and 60+ years (continuous line). The differences between the groups are significant ($p < 0.001$) (after [65])*

6.3 Fast & facile feedback in intranet-working

The traditional methods for obtaining feedback from users during the product development phase have been face-to-face interviews, on-site observations, questionnaires, and laboratory tests. These are both costly and time-consuming to carry out.

Gustafsson [71] of Ericsson Development, Sweden, points out that the traditional methods, with reply cards, or questionnaires do not work well and that the

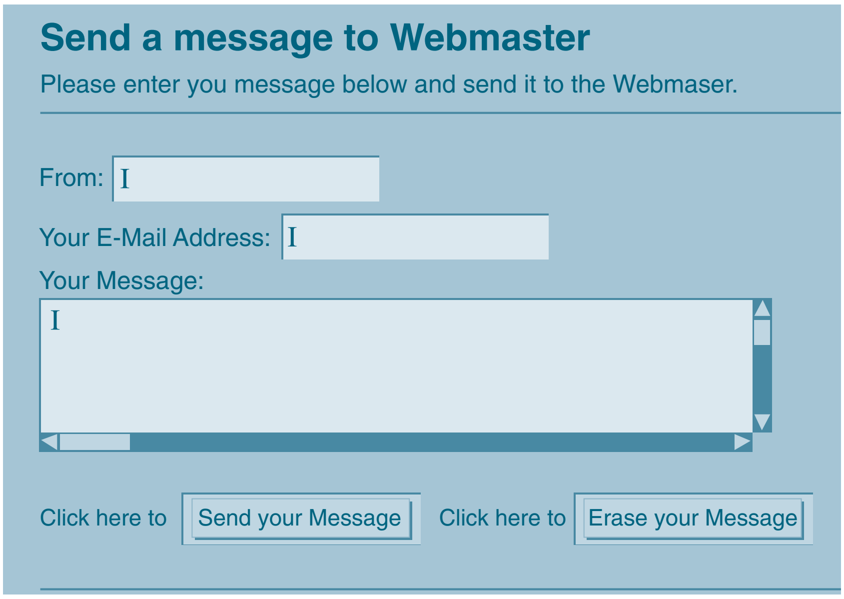


Figure 20 An example of a simple, well designed, web-based feedback tool. Informal, candid and impromptu comments can easily and quickly be sent to the developers' newsgroup (after [71])

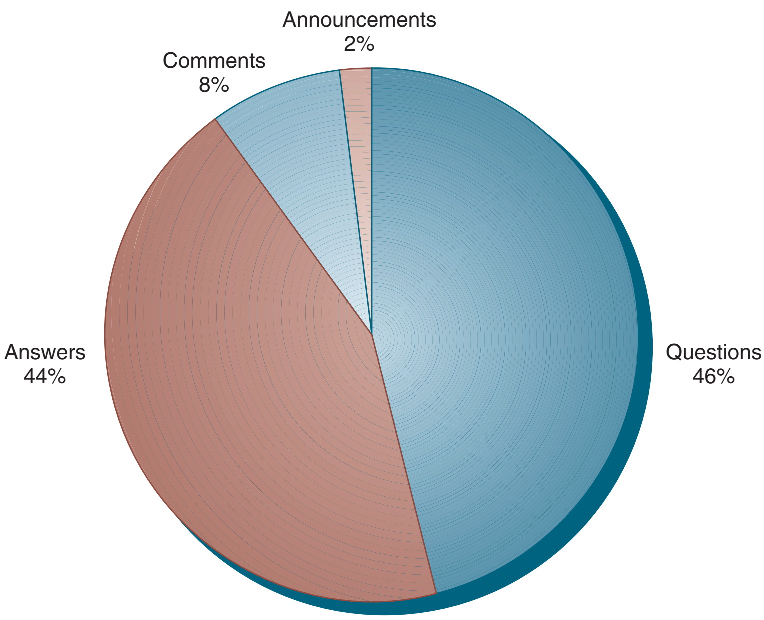


Figure 21 The distribution of news categories (N = 128) received over one year by the designers' newsgroup. Questions (46 %) and answers (44 %) make up the dominating majority of inputs (after [71])

most obvious feedback systems, intranet or internet e-mail, have not been properly utilised for this purpose. He proposes a system called *Mail-order Help,* which is installed with the software to be assessed.

The *Mail-order Help* is a function that was originally conceived as a "What-the-heck?!" dedicated function key. The idea was to be literally only one keystroke away from feedback help, but this idea was temporarily abandoned since users could come to expect that this function key would work with all applications.

Instead, the *Mail-order Help* function was implemented as a help menu option. When selected a simple form is presented asking the user to fill in any questions or comments and then simply select the <Send your message> screen button (Figure 20) to transmit the user feedback to the appropriate recipient (no e-mail addresses needed, no alternative check boxes to be selected, etc.). In order to provide prompt replies, the e-mails were received by an e-mail alias and redistributed to the members of the development team. The team member who replied also had the reply distributed automatically to the other members of the team so that they could know how the feedback had been dealt with.

When a particular *Mail-order Help* feedback was considered to be of more general interest, the question or comment was added, along with the reply, to a newsgroup in order to make the information available to the whole development group. However, the name of the person who made the feedback was deleted and replaced by "Someone just asked …", permitting users anonymity to ask 'silly' questions (Figure 21). Depending on the confidentiality of the development project, feedback and answers can be shared in a larger newsgroup within the organisation, or even outside over Internet.

Gustafsson found that substantial savings could be realised when the help menu option and newsgroup were used in the development work on an operation & maintenance product (OMGR Object Manager for use with Ericsson's ATM Broadband System).

In conclusion, Gustafsson sums up his findings with the e-mail method as easy to use, cost-efficient and much appreciated by the users. Little effort was required to deal with the *Mail-order Help* and the newsgroup, and the benefits were:

- Suggested changes got early feedback from users and designers
- Feedback from users could result in fast improvements.

Another valuable feature of a newsgroup is that it acts as its own tutorial: Beginners may start by reading what the more seasoned developers have written, and when they feel more confident they may start submitting their own replies and comments. As an alternative to more traditional methods, the use of the intranet and Internet for e-mail and newsgroups may have much to recommend it.

7 New services, user requirements and marketing

The proliferation of new services and equipment is accelerating beyond all expectations as the result of very aggressive marketing, while knowledge of user requirements and reactions to the new services are lagging behind. Some new products and services like GSM cellular mobile telephones and the Internet have been taken up despite often inferior user interfaces. This may be because it was felt that the products fulfilled some important needs and functions, i.e. the perceived status, freedom and mobility offered by the mobile telephones, or by the huge amount of information available on the Internet and the relatively low price of accessing it.

Still, it is a fact that many people feel estranged by the new sophisticated techniques, often requiring computer literacy or advanced knowledge of complex user interfaces to be utilised. Even the lovely telephone system is now pervaded by interactive speech-based services where the users have to dial their way through several levels of menus even for very simple tasks such as calling a taxi or ordering a cinema ticket.

7.1 User expectations of metering unit displays and cordless telephones

In a domestic setting, does having the actual cost of using the telephone continuously displayed have any positive or negative effects on the use of the telephone, and does the use of cordless telephones have any positive or negative effects on the time of telephone use? Or in other words, is there a price of being social?

These questions were addressed as part of a larger project by Ling & Hareland [72] from Telenor Research and Development, Norway, in a survey conducted among 200 residents in the Bryne area in southern Norway. 100 residents were provided with a metering unit display (MUD) that continually displays the cumulative costs of an outgoing call to the caller. 100 other residents were provided with a cordless telephone. The hypothesis was that providing technology that displays the consumption of metering units or that allows higher availability and making calls in more comfort would in some way influence the telephone use.

The results should be a quantification of a series of qualitative questions regarding the respondents' immediate impressions of the two types of telephone equipment. Before being provided with the new equipment, each of the two respondent groups had to give their immediate opinions as to the positive and negative aspects of the metering unit display and the cordless telephone, respectively.

The authors' findings show that these technologies stimulate a mirror pattern of associations. The negative aspects of cordless telephones, i.e. that increased comfort and availability would lead to increased use of the telephone, increasing costs, were mirrored by the positive aspects of the metering unit display, i.e. that making the cost more prominent would lead to less use of the telephone, reducing costs. And conversely, the positive aspects of cordless telephones, i.e. that increased comfort and availability would lead to increased use of the telephone and more social interaction, was mirrored by the negative aspects of the metering unit display, i.e. that showing the cost of using the telephone more prominently would lead to reduced use of the telephone and less social interaction.

The users' metaphors for telephone use also have a gendered difference; the idea of the telephone as a utilitarian device, mainly maintained by men, or the telephone as a social device, mainly maintained by women.

The fundamental tone in the comments reported by Ling & Hareland is a deeply held notion that the telephone use is expensive. This finding is quite certainly due to the fact that up to the early 80s the unit cost of making a call in Norway was one of the highest in Europe, and telephones in Norway were, in effect, rationed (people had to wait up to two years to receive a telephone subscription, and children actually took over their parents' telephone subscription when parents died and did not report the change to the operator).

Ling & Hareland divide the comments received in the study of *metering unit displays* into two categories, positive and negative. The positive comments revolved mainly around three themes; monitoring costs, saving money, and saving time. The negative comments centred around the notion that increased cost awareness inhibits spontaneous telephone use which affects the social nature of telephone conversations – it makes one think twice before making an 'unnecessary' (i.e. social) call. This kind of comments underscore the importance of the telephone for holding friends and relatives together.

These findings conflict with the basic hypothesis of the study, i.e. that accurate information on the costs of telephone use would result in more telephone use since most users actually overestimate the cost of telephone conversations.

The comments received on *cordless telephones* were also divided into positive and negative. The positive responses fell mainly into two broad groups, dealing with mobility and with convenience. The most obvious comments were that cordless telephones allow users a much broader range of mobility, both indoors and out of doors, which fixed telephones can not offer. Typical comments were that it allowed more privacy, parallel tasks, and not least, less risk of missing incoming calls; i.e. avoiding having to rush to answer a ringing telephone. While parallel tasks are tolerated, they are often seen as breach of good form. Social norms dictate that a certain amount of attention should be given to the conversation. If background noises (e.g. children bathing or kitchen sounds) become too intrusive, the caller may come to question the attention given to the conversation and the legitimacy of the interaction.

The negative comments on cordless telephones fell into two groups; those dealing with user behaviour and those dealing with the characteristics of the cordless telephone. The issue that received most concern was the fear that the convenience of using cordless telephones would lead to increased telephone use. According to the authors, this is similar to the reasons given by the Amish people in the USA. Although they have chosen to accept the telephone, they do not encourage its use for 'idle' gossip on moral grounds and try to restrict its use by not having it in their homes, but placing it inconveniently in an out building. Although the Norwegian rationale was economic rather than moral, the notion that greater comfort provided by a cordless telephone would lead to more calls was the same. Some other practical issues were concerns that the cordless handset would be easy to misplace and that users would become dependent on the device, and, not least, that short battery life would be a problem. These findings provide insights into the folk definition of telephony, and especially the tension between the consideration for cost and sociability is set into relief. It is against this background that new technologies must be domesticated.

7.3 Societal trends and the telecom markets

The ongoing transition from an industrial to an information society will ultimately affect most areas of life. The convergence of media, computers and telecommunications made possible by the new ICT technologies is at the core of this development. These changes will have significant effects on work, leisure, consumer behaviour and social life. On the one hand, this will lead to more integrated systems, on the other hand it will lead to world-wide networking. As a result of these changes and the opening up of the European telecommunications market in 1998, operators are faced with the challenge of defending their existing markets and to develop new market sectors by introducing innovative technologies and services.

To meet these challenges, Schyguda, [73] from the Technology Centre, Deutsche Telekom, Germany, studied the influences of various trends. Using the WWW-Online-Expert forum set up by Deutsche Telekom, societal, economic and individual trends in Germany and Europe were extrapolated to gather information about future markets. For such work the author proposes interlinking four core skills:

- Market development potential:
  - Socio-economic environment of product planning
  - Initiation of markets
  - Issues of user acceptance
  - ICT-technologies and user behaviour
- Methodology:
  - Analyses of the future
  - Workshops
- Technology monitoring:
  - Early recognition
  - Discovering and assessing innovations
  - Transfer innovations into development scenarios
- Co-operation:
  - Group management
  - R&D strategy
  - Select experts.

The needs and uses of telecommunications depend upon a number of interdependent socio-cultural, political, economic and technological factors. Aggregations of general influences, developments and trends pointing in the same direction Schyguda calls *megatrends.* To illustrate the object of the study he gives an example.

A most significant demographic megatrend is that the percentage of senior citizens in the populations of the industrialised countries will grow rapidly during the next few years. There are currently 16 million elderly people in the Federal Republic of Germany, which comes to about 20 % of the total population, and this is expected to rise to at least 25 % by the end of this century. The characteristics of the target group can be used as a basis for inferring what needs must be taken into consideration by the providers of telecommunication services and equipment. In the case of senior citizens, this would cover such characteristics as reduced mobility, high demands on quality, strong family ties, declining fitness and more available time. Other socio-cultural megatrends would be the ‘generation X syndrome’ or the ‘Nintendo kids’ craze.

In conclusion, Schyguda finds that the methods of ‘trend gurus’ such as Naisbit, Popcorn, Horx & Gerken are often vague on methodology, use ambiguous terminology, lack substantiation and source references, contain imprecise information on trend diffusion over time, and contain too many inconsistencies. Therefore, it is important to develop new techniques.

7.4 Matching telecom services with user needs

A very central question is; are the structures of telecommunication services based on an overall view of the user communication needs? This question is central in the work reported by Rantaaho & Leppinen [74], Helsinki University of Technology, Finland. They have developed a classification of extant telecommunication services that is based on user communication task analysis. As a case they first analysed the communication needs and problems of two mobile care-work organisations. They then analysed existing telecommunication services and grouped them to meet the specific needs of the organisations. This was also carried out for single users.

Mobile care workers calling at the clients’ houses were observed during work. 20 workers were interviewed about their communication with their organisations and participated in simulated communication with a group of six home care workers. The data were then used as the basis for grouping existing telecommunication services according to the needs of the home care organisations. A similar, but more general method was used to analyse the communication needs of 15 single individuals from all walks of life and tabulate the findings.

The authors find that the current trends of *service personalisation* and *service integration* would provide good solutions for service selection problems and message chaining problems for most of the respondents. As an example the authors mention that a message left using a particular service, e.g. voice message or e-mail could be picked up by any service independent of the service used for sending the message. In conclusion, the authors warn about mixing up information about user communication needs and marketing information on estimated use of new services.

7.5 Introducing new telecom services

Creating and introducing new telecommunication services and products in a rapidly re-regulated telecommunication market pose some special problems. Svendsen, Stenvold, Folkow & Akselsen [75] from Telenor R&D, Norway, have devised a method for creating and introducing new telecommunication services and products.

The authors first analysed a number of R&D projects for promoting new telecommunication services and products such as tele-medicine, virtual corporations, tele-work, distance education, communication within local public sector, and monitoring the environment.

In analysing the results, the authors found that the activities that produce deliverables leading to project success fell into four categories:

- Creating examples
- Mobilising opinion leaders
- Exposing obstacles
- Acquiring knowledge.

In analysing the effects of projects to create, promote and introduce new telecommunication services and products, the authors highlighted two problems:

- Services are becoming more complex and are more integrated into peoples' lives
- A service must reach some critical mass before being useful.

These problems were addressed by:

- Create real-life examples to reduce complexity for the user
- Mobilise opinion leaders as promoters to reach critical mass
- Disclose any obstacles.

The authors find this method useful for projects involving the following problems:

- Users need help to see how to adapt services to their needs
- Users do not exploit technical possibilities
- Use of a service depends on others using the service.

Thus, there is a clear need to address these specific telecommunications problems to improve the uptake of new services and technologies.

8 Convergence – integration of telephony, computing and the Internet

The most conspicuous current trend in telecommunications is the integration of telephony and computers – i.e. the convergence of different media. This merging of different media will create a multitude of conceptually new possibilities, possibilities that until recently belonged to the realm of science fiction. Combining telephony and video technology gave us the videotelephone, using the telephone network with computers opened the way for Internet, e-mail and the WWW, integrating the mobile telephone with lap-top computers has given us the portable office, the list may be extended. Integrating microprocessors into different equipment, such as PBXes and terminals, will give us many new services and features that were quite unthinkable until recently. Intelligence in the network will be able to keep track of all subscribers and route calls according to urgency and personal priorities and agendas. Hopefully, such systems may one day serve us as a loyal, but now nearly extinct, secretary or butler – for some reason robots in science fiction films all seem to speak the distinct and correct English of a well trained butler.

The problem today is that most of the sophisticated new possibilities are underused; though many people agreed that they wanted them when they were asked. Often only one or two new products seem to fill a need and are adopted by any number of users, often despite bad usability (cf. the early video recorders), while the majority of new features are seldom, if ever used. The question, then, is how to make all these wonderful new services and features so simple and intuitive to use that people will actually use them.

This of course is the domain of human factors, and here lies a challenge that is different from the traditional human factors in telecommunications. One important source of knowledge to this field is through studies of what people actually do, what they actually need and how they redefine services and equipment according to the changing ways of life. To take one example: diverting telephone calls to another number has been a feature of PBXes for many years, you use it if you go into another office, but expect an important call. When this service was introduced in the public network, it was given the code: *21*<telephone number># (and #21# to cancel the service). This is not a particularly user-friendly code, mainly because there are a large number of other supplementary service codes to remember. From the start the service was not well known and therefore not much used.

However, burglars in Norway have a *modus operandi* where they phone prospective homes to find out if they are inhabited, especially during the Easter and Christmas holidays. Then the popular notion spread that if you diverted your calls to a neighbour or a friend when you were on holiday, a burglar who calls would be met by an answered phone and mistakenly believe that the home was inhabited and go somewhere else. This use of this supplementary service became so popular, especially during Easter and Christmas holidays, that the capacity was exceeded and many prospective users were met by the message that the service was unavailable.

This is just one example of a simple service, which at the outset had no more potential than many other supplementary services, but that was taken up because of a life-style and concerns about being burgled. We will see many other such success stories in the future, but we will not be able to know beforehand which service will be a hit. One important task will therefore be to monitor the life-style and actual needs of subscribers so new trends can be assessed as soon as they appear. Operators who do not see the trends and react to them in time may lose much revenue.

With the convergence and integration of different services and media, we are going to see some very spectacular new services, but which services will make it in the marketplace and which services will be failures cannot be determined *a priori* – new trends are never easy to predict. Sometimes even the most unlikely products become hits, cf. the Walkman® or the Tamagoshi®, products that only a year before they appeared would be regarded as unthinkable.

Human computer interaction (HCI), graphical user interfaces (GUI), object manipulation interfaces, may ultimately provide services that are very simple to use. This trend is emerging very fast: Why use cryptic codes and ungainly procedures when a graphical user interface and object manipulation would make the service much more user-friendly?

8.1 'Uncountable' smart features that no-one uses

A large number of new 'smart' features in PBXes, to be accessed from terminals by cryptic codes, do not necessarily improve usability or provide efficient and satisfactory telephone services, and are therefore seldom used. Lindgaard [76] probably put this most bluntly in her paper: "Wow – 568 Smart Features on Your PABX. What Really Determines the Uptake of Technology?"

Will adding another package of functions improve the usability of a PBX or a smart terminal if the user requirements for representation of services and interactions are not carefully considered? The answer to this is no, according to Herstad*, Hüttenrauch# & von Niman# [77] from *UNIK, Norway, and #Ericsson Infocom Systems, Sweden. They present findings and conclusions from a user interface development process and usability studies with a PC-based graphical user interface for integrated computer telephony.

Based on a metaphor of desktop application-windows and graphically represented services and following repeated cycles of end-user need studies, usability studies, prototyping and software engineering iterations, a computer telephony integration (CTI) application called *Personal Screen Call*™ was introduced by Ericsson.

The visualisation of available services and options without the need for recalling codes or having to look them up was a clear advantage over traditional systems, and the system met with much user approval. The availability of up-to-date directories, the simplicity to select and click icons to access services was especially well received.

However, there were also some important problems. Computer systems are still not as reliable as the traditional telephone system, and when the system crashes you loose all your fancy telephone options, even though an ongoing conversation may not be disconnected. Also, setting up a user-friendly and practical system can be a problem. Just providing a user with the installation diskettes may not necessarily result in a positive first-time encounter with CTI telephony.

As traditional supplementary services are getting more complicated, telephone numbers are increasing in length and the number of available features and options in the new systems is growing all the time, CTI telephony is certainly going to win in the end. However, there will be an imperative demand that the new systems must be at least as secure and safe as today's telephone systems; secure against crashing, safe against important information being lost. We have come to trust the traditional telephone system to the point that we come to entrust our personal health and safety with the telephone. When this level of security can be achieved and guaranteed in CTI telephony, and is generally accepted by the users, CTI telephony will be the preferred way to communicate. Until then, we will have to live with a system that provides a multitude of smart features that no-one uses because the user interface is just too bothersome.

8.2 Your intelligent agent on the web

In order to fully take advantage of the many advanced new telephony services that soon will be offered, subscribers will be required to enter information into the system computer and update the data associated with their accounts. To achieve this, the subscribers must be able to interact with the system computer that controls the service. This is normally performed over a very restrictive communication channel, usually making use of touch-tone (DTMF) or voice input and auditory feedback. This restrictive channel is highly unsatisfactory to most users, and this creates a great hindrance to the acceptance of new advanced services. What, then, is the preferred manner for subscribers to interact with the service computers of these advanced systems?

Mané & Rabin [78] of AT&T, Holmdel, NJ, USA, have addressed this question. They also encountered this interface obstacle when designing the user interface for a telecommunication 'intelligent agent' service. To by-pass the restrictions imposed by the telephone-based interface, they chose instead to develop a graphical user interface (GUI) that is available to the subscribers over the Internet. The authors' main goal was to explore how 'intelligent agents' can effectively be used to offer subscribers new features and services in the telecommunications field. To explore this field they developed a *personal assistant for telecommunications* (PAT) that offered the functionality resembling the services normally provided by a telephone receptionist. To be a truly personal assistant, the service must be customised so that the 'agent' can tailor its services to the needs of each subscriber. The authors also chose to forego machine learning, which is usually an important attribute of intelligent agents, and instead make use of user-defined rules that rely on the subscriber supplying the information. This information ranged from simple operations, such as furnishing a telephone number, to complex specification of the different options to be presented to a caller, or on how calls should be routed according to personal needs and wishes, requiring a solid grasp of how the system worked. Because of the amount and complexity of the date to be entered by the subscribers, the authors chose to employ a graphical user interface accessible over the Internet rather than the common interactive telephone interface.

To evaluate the intelligent agent the authors recruited eleven computer-literate employees with no prior knowledge about the agent to fill in a service registration form presented via a Netscape Navigator® browser running on a networked PC. The evaluators also filled in a short questionnaire and discussed it with the experimenter.

A registration took on average 9.7 minutes to complete. All except one subject found this acceptable. Although the tested interface supported only the *administration* of the service, the GUI interface may equally well be used to *control* services. The evaluators revealed several flaws in the system; one was that many users actually forgot to submit the data they had entered to the system because the *Submit* button was not very prominent and easily overlooked. The authors' most interesting finding, though, was that 91 % of the evaluators actually preferred to use the web to using pen-and-paper to register for the service.

Mané & Rabin find that they have learned some valuable lessons from their study: The *KISS* principle (Keep it simple, stupid) is probably the single most important factor when designing a user interface. One must remember that for a system to be usable for most subscribers it must also be usable for the less bright users. Half of all people are less than average intelligent (this is an inevitable effect of the definition of the intelligence scale) but a factor that is easily overlooked since most engineers, designers and programmers are academics, and academics are predominantly recruited from this upper half of the intelligence distribution. Engineers and designers usually design for themselves and their peers, which means that large groups of users have to cope with systems that may often be too difficult for them to understand and use. Engineers and designers may find it difficult to put themselves into the position of less intelligent users, but by following the KISS principle they may still come up with usable designs.

Take the users' perspective is another lesson, but as just mentioned, this may be difficult for some engineers and designers. The problem is that designers have too much knowledge about the system they are designing to place themselves apart from it and see it with the eyes of the user. Nonetheless, this is an important exercise to perform whenever possible and should not be too difficult with some practice.

Take advantage of the visual display. A visual display offers the use of a graphical user interface, which is much more effective when dealing with complex information than a monotonous reading of an auditory menu. For simple interactions (e.g. replying 'yes' or 'no', or making a simple choice between two options) an auditory prompt may be just as, or more efficient.

Design for incremental learning. The user rarely needs to know all about a system to use it. It is far better if the user can start to use the system with only some initial knowledge, and then learn as he/she goes and new features are taken up. Nothing is so discouraging as to have to read a thick user manual before being able to use a new service. Give the user a simple start-up instruction and build on common knowledge.

Present only task-critical information. Only information that is really necessary for performing the task at hand should be presented, making the task simpler and faster to complete. This principle also applies to help information – it is most frustrating getting the whole instruction manual displayed on the screen when asking for help.

Ease of communication and ease of software distribution in web design. As shown above by Gustafsson [71], the web can actually be a most efficient vehicle for user interface design through ease of communication and ease of software distribution. Using the web actively provides the designer a number of advantages, not least the fast and easy environment for collaboration, both with peer designers and with evaluating users.

The authors conclude that the use of graphical user interface was most enthusiastically received by the users and could also be applied to many other telephone services besides the 'intelligent agent', both alone and in combination with interactive voice interfaces.

8.3 Is the 'surfing' metaphor bad for novices?

The *metaphor,* i.e. a resemblance, likeness or similarity, is a powerful tool for conveying information on something completely novel by using well known similarities. Not least in modern information technology we use a number of metaphors, well-known examples are the *memory* metaphor, the *desktop* metaphor, and the *window* metaphor. The World Wide Web is an Internet service for accessing information and services that is growing fast in popularity, not least due to its graphical user interface (GUI) that allows users to utilise the web without knowing the underlying computer network architecture. It is thus possible to begin 'surfing' (another metaphor) the web by use of any of the many web browsers using a graphical user interface without any prior training or guidance.

Does the 'surfing' metaphor actually communicate a correct mental user model of what is going on? What mental models are created by the users from the often arbitrarily chosen metaphors? Is a well-chosen metaphor conductive to presenting the Internet in a manner that does not contradict the functioning and system architecture, or is any correspondence between the system and the metaphor superfluous? These questions are addressed by Ranta-aho [79] from Helsinki University of Technology, Finland, who in an experiment investigated the way users perceive systems and create mental models.

Learning to use a computer system (or any other complex system) usually involves constructing mental representations of some aspects of the system. This mental model is then utilised for predicting the system's behaviour. A model need not be a 'correct' representation of what actually is going on in a system (few are) to be usable, but a model should not contain elements that may lead to incorrect actions. A well-chosen metaphor may induce the formation of a constructive mental model, but few metaphors have a wide enough scope to encompass all the possibilities, and most metaphors quickly run into trouble when stretched beyond their original meanings. Many metaphors emanate arbitrarily, often because of some conspicuous, but insignificant likeness, and may therefore convey false assumptions to the user of how the system (*viz.* the Internet) functions. According to the author, many such false assumptions are acquired through public media.

Ranta-aho interviewed ten novice users who were reasonably computer literate (5 male and 5 female, aged 20–51 years) before their first experience in using the WWW, focusing on their ideas of what was happening when they used the web. The interviewees only knew about the web from newspapers, TV and what they had heard from friends. They were also given a seven question questionnaire (see Table 3). In addition twenty persons participating in a WWW course were interviewed before and after the course on their ideas of what happens during the

Table 3 Problem solving questions used in the interviews

1) While using the WWW, when you click with the mouse onto a link, what happens?
2) Is it possible to get some kind of virus using the WWW?
3) You are currently reading on-line a web page of a foreign university. Suddenly, someone stumbles over your computer's electricity cable and your computer gets switched off. Will you stay as user in the computer of the foreign university?
4) You are reading some web pages. Suddenly, your network connection breaks. Is it possible for you to still see some of the pages you just read?
5) You are reading a web page, when the author of the page updates it at the same time. Can you see the changes immediately?
6) You are reading a web page with a form to fill in, and you start filling in the form. At the same time, another person starts reading the same page from his computer. Will he see your entries?
7) You read a web page of a company. Where is the actual file located, that you are reading?

use of the web. They were also issued the questionnaire, both before and after the course.

The novice users who had not used the web before as well as the people taking the WWW course were able to explain their notions about the functioning of the web with some simple model. These models fell into three basic categories, which were characterised by different levels of understanding what was really going on:

1 ***Fetch files to my computer*-explanations** contained file movement related terms such as files being copied to own hard disk. Typical explanations were "The browser asks other computers to send files to user's computer", "The files shown are copies of the original files, copied to the own hard disk".
2 ***Surf around*-explanations** were identified by usage of own movement related terms that imply that the web browser temporarily logs onto the other computer and reads the web page files. Typical explanations were "I go to that page", "I am in the other computer".
3 ***Print onto my screen*-explanations** fall between the first two and are characterised by wordings like "from the other computer, my browser shows the file on the screen". They do not suggest "being logged as user" to the other computer nor do they suggest "files being transferred to their own computer", but just "being shown on its screen".

The interviewees' replies fell into three groups:

1 Offered 'Fetch files'-explanations and correct answers to all seven questions (2 of the 10 novices, 4 of the 20 WWW course participants).
2 Offered the 'Surf' or 'Print'-explanations and had one or more wrong answers to the seven questions (8 of the 10 novices, 12 of the 20 WWW group).
3 Also gave the 'Surf' or 'Print'-explanations, but actually knew the right answers to the questions. When prompted they explained "he really does not go anywhere but his computer fetches the file. But he is used to talk about it like this" (0 of the 10 novices, 4 of the WWW group).

The author also notes that:

- All subjects, also those who had never used the web, were able to give some form of explanation about the web.
- Some subjects used the 'Surf' or 'Print' metaphors even though they actually had more exact knowledge about the browser/web interaction (Group 3).
- Some subjects used the 'Surf' or 'Print' metaphors, but could not always predict correctly the behaviour of the browser (Group 2).
- Training in the WWW course had no effect on the explanation model subjects used.
- The quality of the answers from the novice group and the WWW course group were very similar.
- Only one of four wrong answers to questions 2–7 by participants of the WWW course were correctly answered after the course.

Ranta-aho's results suggest that a mental model of the web may be created even without any practical experience with the web. An inaccurate mental model can be very resistant to correct information. Those who have a correct mental model tend not to be able to predict important aspects of the web browser behaviour more correctly than those who hold an incorrect model are.

In conclusion Ranta-aho stresses the importance of the user interface reflecting the users' task instead of the system's properties [80]. This can lead to interesting results if some users generate system architecture models based on task level metaphors. Even if users do not need to understand technical details of computer networks, they may soon have to understand the basics of computer networking to be able to work correctly and safely on the web. Such information can also be presented on the user task level by using "I sit at my computer and all information flows to me"-metaphor instead of "I surf around the world"-metaphor.

9 Conclusions

Several of the new solutions reviewed above may have significant effects on our future use of telecommunications and on the way we will interact with the new systems and services. Gone are soon the days of the simple telephone number dialling, and instead voice recognition, graphical user interfaces, the Internet and intelligent agents with the ability to adapt to our needs will be our future communication intermediaries. Today's ubiquitous telephone numbers may soon be replaced by address systems based on the subscriber's name, place of work or other easily remembered distinguishing characteristic; the ungainly digit strings required to access complex services may soon be superseded by object-oriented graphic user interfaces; the interactive voice-based services requiring keypresses in response to prompts, will soon give way to intelligent agents that discern the subscriber's needs and wishes; and mobile services have already started to free subscribers from fixed communication sites, a development that will only spread world-wide as low-orbit satellites or stabilised air-ship-borne transponders are introduced.

Internet and the World Wide Web have already made great inroads into traditional telecommunication services, and increasingly more of the new services will be based on information technologies (IT) rather than on the traditional public switched telephone network (PSTN). This makes it more important than ever before to provide simple, easy-to-learn user interfaces that will not exclude users, especially people who are elderly or impaired. Actually, IT-based services are inherently easier to configure to individual user needs than the traditional telephone services.

The cost/benefit of improved usability of telecommunication services has not been extensively documented, except for internal operator procedures, where substantial savings have been achieved by rather small measures. The same potential savings are possible by paying attention to the interface for the end-user. This issue has now been placed firmly on the agenda, and there is now an increasing awareness among operators and developers of the potential benefits and savings that improved usability can give.

Finally, standardisation has been mentioned by some of the authors. In the days of the state-owned telephone administration monopolies, the end-user had no way of influencing the design of telecom services or the pricing policies, except very indirectly through general elections. Standardisation of telecommunications in those days was carried out in Europe by CEPT (*Conférence Européenne des Administration des Postes et Télécommunications)* and globally by CCITT

(Comité Consultatif International Télégraphique et Téléphonique, now ITU-T) to achieve trans-border and trans-service interoperability, but users and consumers had virtually no way of influencing the standards adopted. When the open European telecommunication market was created, the Commission of the European Communities (CEC) set up the *European Telecommunication Standards Institute* (ETSI) to create European regional telecommunication standards. ETSI, however, was not only open to the telecom operators and the regulatory authorities, but was also open to industry, universities, research organisations, and even to private organisations and statute administrations, including user groups and consumers' organisations.

With the increasing deregulation of the telecommunications sector and the growing privatisation of the previously state-controlled operators, the main emphasis of the operators has shifted more towards making profit than to provide proper services. One way for the operators to minimise overheads has been to decrease participation in standardisation activities. However, new actors are now rapidly filling this void. Industry has already been mentioned, as have universities and research groups, but the most recent actors are the user groups and consumers' organisations. ETSI has even established its own User Group to supervise standardisation and see that the users' perspective is maintained in the new standards. It may turn out to be a short-sighted, ill-advised and rather costly decision by the telecom operators to reduce participation in standardisation and leave the arena open to the special interests of user groups and consumer's organisations. Actually, the interests of the telecom operators and the user organisations may often coincide, but the technical solutions that could be proposed by the user organisations as part of the new standards may not always be optimal for the efficient and economic provision of telecommunication services. It is therefore important that the operators find a strategy for participation in standardisation work in close co-operation with the user representatives.

10 References

1 Nordby, K (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997. (ISBN 82-994236-0-0.)

2 ETSI ER 145. *Human Factors : European harmonization of network generated tones. Part 1: A review and recommendations.* Sophia Antipolis, ETSI, 1997.

3 ETSI ER 145. *Human Factors : European harmonization of network generated tones. Part 2: Listing and analysis of European, World and standarized tones.* Sophia Antipolis, ETSI, 1997.

4 Gagliardi, D. *Report on the audible tones in the telephone service in the EC countries.* Report from Contract 48159, DG XIII. Brussels, Commission of the European Communities, April 1993.

5 ETSI ETR 170. *Generic user control procedures for telecommunications terminals and services.* Sophia Antipolis, ETSI, 1995.

6 Anderson, D M, Ferris, A, Mellors, W. Human factors in ETSI : use of tones in public telephone networks. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 3–14. (ISBN 82-994236-0-0.)

7 Helmreich, R. Ergonomic user interfaces on telephones – are they worth it? In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 15–20. (ISBN: 82-994236-0-0.)

8 Dieterich, H, Steiner, D. Intelligent agents as appointment managers living in public networks. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 23–30. (ISBN 82-994236-0-0.)

9 FIPA. Foundation for Intelligent physical agents" [WWW site] *http://www.cselt.stet.it/fipa* (visited February 1997).

10 Goldstein, M. A common functional interface for all Types of voice response services. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 33–40. (ISBN 82-994236-0-0.)

11 Schmidt, M et al. User perspectives on the security of access data, Operator handover procedures and 'insult rate' for speaker verification in automated telephone services. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97)*, Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 41–48. (ISBN 82-994236-0-0.)

12 Charry, M C, Zuñiga, M C, Mejia, J J. An experience of a telecom CTI system. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 49–56. (ISBN 82-994236-0-0.)

13 Hamnes, K et al. Dialogue design for numeric paging using automatic speech recognition. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97)*, Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 57–64. (ISBN 82-994236-0-0.)

14 Fay, D. User acceptance of automatic speech recognition in telephone services. *Proceedings of the International Conference on Spoken Language Processing, Yokohama,* 1994, 1303–1306.

15 Karis, D. Speech recognition systems: Performance, preference, and design. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97)*, Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 65–72. (ISBN 82-994236-0-0.)

16 Kosko, B. Hidden patterns in combined and adaptive knowledge networks. *International Journal of Approximate Reasoning*, 2, 1988, 377–393.

17 Kosko, B. *Neural Networks and Fuzzy Systems.* New York, Prentice-Hall, 1995.

18 Zajicek, M et al. Answering vague enquiries in computer aided telephony. In: K. Nordby (ed.). Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97), Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 75–82. (ISBN 82-994236-0-0.)

19 ITU-T. *Modulated Noise Reference Unit (MNRU).* Recommendation P.81, Blue Book, 5, 1989, 198–203.

20 ITU-T. *An On-going Series of Subjective Experiments to Assess Speech Output from Text-to-speech Systems.* COM-12-18-E, June 1993.

21 ITU-T. *Report of the First Meeting of Study Group 12 (Geneva, 10–19 May 1993), Part B : Draft ITU-T Recommendation P.85 : A Method for Subjective Performance Assessment of the Quality of Speech Voice Output Devices.* COM-12-R-6-E, June 1993.

22 Watanabe, T. TFW-PM reference for subjectively assessing quality of synthetic speech. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97)*, Oslo, May 12–16, 1997. Oslo, Norwegian Telecommunications Authority, 1997, 83–90 (ISBN 82-994236-0-0.)

23 Hick, W E. On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, 4, 1952, 11–24.

24 Davis, R, Moray, N, Triesman, A. Imitative responses and the rate of gain of information. *Quarterly Journal of Experimental Psychology,* 13, 1961, 78–79.

25 Thomas, B. Choice reaction times in voice applications. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97)*, Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 91–98. (ISBN 82-994236-0-0.)

26 Mowbray, G H. Choice reaction times for skilled responses. *Quarterly Journal of Experimental Psychology*, 12, 1960, 193–202.

27 Goldstein, M. Classification of methods used for assessment of text-to-speech systems according to the demands placed on the listener. *Speech Communication*, 16 (3), 1995, 225–244.

28 Delogu, C, Paoloni, A. Evaluation of speech synthesis systems for telecommunication services. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 99–106. (ISBN 82-994236-0-0.)

29 Walker, H W, Sproull, L, Subramani, R. Using a human face in an interface. In: *Proceedings of CHI'94, Human Factors in Computing Systems.* Boston, CHI, 1994, 65–91.

30 Taylor, I C et al. User trust in talking heads as providers of information in multimedia interfaces. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97)*, Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 109–116. (ISBN 82-994236-0-0.)

31 McGurk, H, MacDonald, J. Hearing lips and seeing voice. *Nature,* 264, 1976, 746–748.

32 Sekiyama, K, Tohkura, Y. McGurk effect in non-English listeners: few visual effects for Japanese subjects hearing Japanese syllables of high auditory intelligibility. *Journal of Acoustic Society of America*, 90 (4), 1991, Pt 1.

33 Contin, L et al. Methodological studies on subjective evaluation of audio-visual quality. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 117–124. (ISBN 82-994236-0-0.)

34 Contin, L et al. Impact of video coding degradation on emotion recognition. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 125–131. (ISBN 82-994236-0-0.)

35 Tetzchner, S von. The use of 64 Kbit/s videophones for sign language. In: S. von Tetzchner (ed.). *Issues in Telecommunication and Disability : COST 219 Project.* Brussels, Commission of the European Communities, 1994, 327–339. (ISBN 92-826-3128-1.)

36 Kamata, K. Sign language conversation through videophone in ISDN service. *Proceedings of the 2nd European Conference on the Advancement of Rehabilitation Technology (ECART 2),* Stockholm, Sweden, 1993.

37 Kamata, K, Hirama, A. The use of 64kbit/s videophones for hearing impaired people. In: *Proceedings of the 15th International Symposium on Human Factors in Telecommunications (HFT'95),* Melbourne, Australia, 1995.

38 Kamata, K, Hirama, A, Oguro, S. ISDN videotelephones for the people with hearing disabilities. *Proceedings of the 3rd European Conference on the Advancement of Rehabilitation Technology (ECART 3),* Lisbon, Portugal, 1995, 276–278.

39 Kamata, K, Yamashita, M, Hirama, A. Sign language conversation and picture quality of 64Kb/s ISDN videophones. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97)*, Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 133–138. (ISBN 82-994236-0-0.)

40 Fosse, T V, Dolvik, J. Sign language video for hearing impaired, compatible with digital video broadcasting. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 165–166. (ISBN 82-994236-0-0.)

41 Suwita, A, Mühlbach, L. Video-communications as a medium for informal communication? In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97)*, Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 141–147. (ISBN 82-994236-0-0.)

42 Hopf, K, Runde, D. 3D-desktop-displays for multimedia. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 149–156. (ISBN 82-994236-0-0.)

43 Prussog, A, Mühlbach, L, Böcker, M. Telepresence in stereoscopic video-conferencing. In: *Proceedings of the 14th International Symposium on Human Factors in Telecommunications (HFT'93).* Darmstadt, R.v. Deckers Verlag, 1993.

44 Runde, D, Suwita, A. Stereoscopic pointing in desktop applications. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 157–164. (ISBN 82-994236-0-0.)

45 Bias, R G, Mayhew, D J (eds.). *Cost–Justifying Usability.* Boston, Academic Press, 1994. ISBN: 0-12-095810-4.)

46 Petrie, H, Fabrizi, P, Homatas, G. Report on user requirements for the accessibility of graphical user interfaces by blind people. In: *Graphical User Interfaces for Blind Persons : TIDE Pilot Action Project 103*. Report to the CEC. London, Royal National Institute for the Blind, 1993.

47 Stephanidid, C, Savidis, A, Acoumianakis, D. Tools for user interfaces for all. *Proceedings of Second TIDE Congress.* Amsterdam, IOS Press, 1995.

48 Laux, L et al. Designing the World Wide Web for people with disabilities : A user centred design approach. *Proceedings of ASSETS'96, Second Annual ACM Conference on Assistive Technologies.* New York, ACM, 1996.

49 Zajicek, M, Powell, C. Accessing the World Wide Web by telephone. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97)*, Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 177–184. (ISBN 82-994236-0-0.)

50 Beattie, W, Hine, N, Arnott, J. Distributed assistive communication devices for non-speaking users. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 185–190. (ISBN 82-994236-0-0.)

51 Hine, N, Beattie, W, Arnott, J. Study of picture annotation as a means of assisting non-speaking people to use telecommunications services. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97)*, Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 191–198. (ISBN: 82-994236-0-0.)

52 Hine, N. People with special needs and service aspects. In: P. Byerley & S. Connell (eds.). *Integrated Broadband Communications : Views from Race, Usage Aspects.* North-Holland, Elsevie, 1992, 201–230. (ISBN 0-444-89391-1.)

53 Hine, N, Beattie, W, Arnott, J. Factors affecting usability of communications services by people with disabilities. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 201–208. (ISBN 82-994236-0-0.)

54 Van Mele, I et al. Mobile access to WWW services for university students who are motor impaired : a qualitative study on the man/machine interface. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97)*, Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 209–216. (ISBN 82-994236-0-0.)

55 Hellström, G. Quality measurement on video communication for sign language. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 217–224. (ISBN: 82-994236-0-0.)

56 Delvert, J. Verification of quality requirements on video-telephony for sign language : videotelephony and relay service for sign language. A comparison between 4 projects with different technique, bandwidth and services. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 225–231. (ISBN 82-994236-0-0.)

57 Sallnäs, E-L, Hedman, L-R, Virtanen, M T. Assessment of user needs for future work practices. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 267–274. (ISBN 82-994236-0-0.)

58 Böhm, A et al. Use of collaboration tools in distributed research projects now and in future. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 275–284. (ISBN 82-994236-0-0.)

59 Honda, S et al. A virtual office environment based on a shared room realizing awareness space and transmitting awareness information. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 287–295. (ISBN 82-994236-0-0.)

60 Nilles, J. *Making Telecommuting Happen : a guide for telemanagers and telecommuters*. International Thomson Publishing, VNR Computer Library, 1994.

61 Nilles, J. Telecommuting and traffic reduction. *ECTF News*, #6, 1994.

62 Matarazzo, G. Telework in the Municipality of Rome: a case study. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on*

Human Factors in Telecommunications (HFT'97), Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 297–303. (ISBN: 82-994236-0-0.)

63 Papa, F, Spedaletti, S. New broadband technologies in tele-education: A human factors laboratory experiment. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 305–312. (ISBN 82-994236-0-0.)

64 Fleet, G, Talbott, D, Winlow, T. Graphical user interface design for the user delivering value to the customer. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 315–322. (ISBN 82-994236-0-0.)

65 Love, S, Foster, J C, Jack, M A. Assaying and isolating individual differences in automated telephone services. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97)*, Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 323–330. (ISBN 82-994236-0-0.)

66 Egan, D E, Gomez, L M. Assaying, isolating and accommodating individual differences in learning a complex skill. In: *Individual Differences in Cognition,* 2, Academic Press, 1985.

67 McCrae, R R, Costa, P T. Validation of the Five-Factor Model of Personality across instruments and observers. *Journal of Personality and Social Psychology,* 52, 1987, 81–90.

68 Heim, A W. *AH Series, NFER-NELSON,* 1970.

69 Gronwall, D. Paced Auditory Serial Addition Task : A measure of recovery from concussion. *Perceptual and Motor Skills,* 44, 1977, 367–373.

70 Lezack, M D. *Neuropsychological Assessment, 2nd edition.* Oxford, Oxford University Press, 1983, 422–429.

71 Gustafsson, N-E. Intranetworking with users – Fast & facile feedback. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 331–337. (ISBN 82-994236-0-0.)

72 Ling, R, Hareland, M. The cost of being social: User expectations of metering unit displays and cordless telephones. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97)*, Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 365–372. (ISBN: 82-994236-0-0.)

73 Schyguda, G. Societal trends and their impact on telecommunication markets. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97)*, Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 383–390. (ISBN 82-994236-0-0.)

74 Ranta-aho, M, Leppinen, A. Matching telecommunication services with user communication needs. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97)*, Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 401–408. (ISBN 82-994236-0-0.)

75 Svendsen, G et al. A method for creating and introducing new telecom services and products. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97)*, Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 409–413. (ISBN 82-994236-0-0.)

76 Lindgaard, G. Wow : 568 smart features on your PABX : What really determines the uptake of technology? In: *Proceedings of the 14th International Symposium on Human Factors in Telecommunications (HFT'95).* Darmstadt, R.v. Deckers Verlag, 1993.

77 Herstad, J, Hüttenrauch, H, von Niman, B. Now 569 Smart features on your PBX: Increasing usability through Computer Telephony Integration …? In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 427–434. (ISBN: 82-994236-0-0.)

78 Mané, A, Rabin, M. Meet your communication intelligent agent on the World Wide Web. In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 435–441. (ISBN 82-994236-0-0.)

79 Ranta-aho, M. WWW and the surfing metaphor: harmful for the novice user? In: K. Nordby (ed.). *Proceedings of the 16th International Symposium on Human Factors in Telecommunications (HFT'97),* Oslo, May 12–16, 1997. Oslo, Telenor, 1997, 443–450. (ISBN: 82-994236-0-0.)

80 Nielsen, J. *Usability engineering.* NJ, Academic Press, 1993.

Knut Nordby is Senior Research Psychologist at Telenor R&D, Kjeller, where he has worked since 1985. He is Chairman of the ETSI Technical Committee Human Factors (TC HF), of which he was a founding member. He is member of the Permanent Steering Committee of Human Factors in Telecommunications (HFT) and Adjunct Professor teaching Man-Machine Interaction at UNIK, Center for Technology at Kjeller, University of Oslo.

e-mail: knut.nordby@fou.telenor.no

Feedback tones in public telephone networks: Human Factors work at ETSI

DONALD M. ANDERSON

A problem for users?

All telephone users are familiar with feedback from the network – they lift the handset, hear what they know as a 'dialling' tone, dial, then hear a 'ringing' tone. Or do they? Is it a ringing tone or a 'congestion' tone, or something else they do not recognise? The effect may be the same, if they are not connected with speech, they will probably hang up and try later. Users mostly do not understand the meaning of all the tones they may hear. This problem is much worse when dialling to another country or when visiting another country and trying to call home. Even worse, in a payphone they may lose any money they may have inserted, because of such misunderstanding.

A previous generation of users simply lifted the phone receiver off the hook and a human operator asked you for the number you wish to call, perhaps the ultimate form of feedback. The operator would even tell you that the called party was busy or not replying. Then around the 1920s, automation allowed you to dial your own number, so feedback by tones was introduced. Such a simple user procedure may be represented in a classic 'indicate-control-indicate' model, as illustrated in Figure 1.

Under normal conditions, a user no longer had any need for operator intervention, except for 'trunk calls'. This was good human factors practice applied long before the expression became recognised, but at that time, only a limited number of tone codes was necessary. The situation worsened when subscribers became able to dial their own calls into networks overseas, who had adopted different 'foreign' tones. As services proliferated and became more automated for direct access by a subscriber utilising a telephone dial, so even more information was, of necessity, provided in the network for presentation to the user as tones.

Today, in some 15 countries in the EU, and 174 world-wide, some 228 different tones are defined as in use in Europe alone. In addition to feedback to the human user, audible tones may also be used by connected terminals for automatic recognition. For humans, recognition of the meaning of some tones is today largely a matter of guesswork and luck for the majority of users. An example of early concern for the user, from nearly 30 years ago is given by Karlin [1], when the CCITT discussed a contribution based on experimental evidence from five different countries, which showed the degree of confusion that then existed between identification of 'ringing' and 'busy' telephone signals.

The ITU-T (formerly CCITT), was from early on aware of the potential problem that differences between national network tones would cause for users, and by promulgating 'Recommendations', tried to standardise the parameters for audible tones for use in public networks, both by describing the *function* represented by the tone code, such as *'dial', 'ring',* or *'busy'*, and also aspects such as audio frequency, cadence and other signal parameters to ensure signal compatibility between networks, but not necessarily user understanding.

The EU in 1993 recognised that the proliferation of tones caused by continuing expansion of telephone networks and services world-wide had led to potential for confusion and error by the users and in a review carried out by Gagliardi [2] he proposed that standardisation and harmonisation of network tones between EU and European countries should take place. This paper describes the aims and methods used to review the problem and arrive at possible proposals for change. The work was undertaken by a team of human factors specialists working to an EU Mandate, under contract to the European Telecommunications Standards Institute (ETSI).

Review of the problems

A number of specific tasks was proposed in the EU Mandate. These included:

- a review of the research and reports of the Human Factors of service tones
- a review of the standards originating from the ITU-T, CEPT and ETSI
- a review of the technical aspects of tone production in networks, and
- a review of actual national regulations in force in the countries of the European Union.

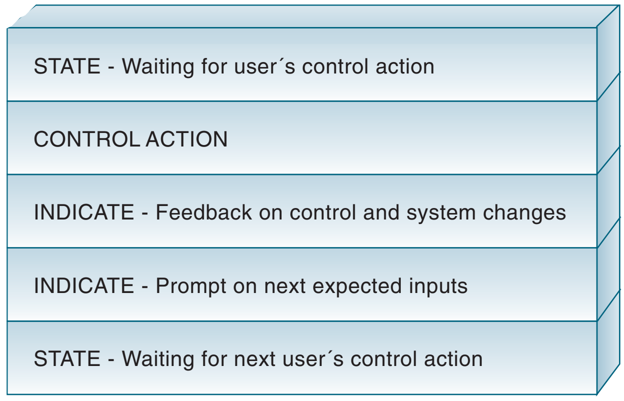


Figure 1 Simple model of user procedure

The complete activity undertaken by the team appears in a two-part Technical Report from ETSI ([3] and [4]). The second part contains what is probably the most comprehensive analysis and up-to-date listing of tones used in the EU, in the countries of the ETSI members and in other countries world-wide published to date. The analysis was of special interest because it revealed that coding was mainly by frequency and especially by cadence. Two examples of tables from the analysis of European tones are given in Annex 1.

Six tones were defined in the Mandate and specifically suggested as priority candidates for standardisation, including:

- ***dial, ringing, busy, special information, call waiting and pay tones.***

A further set of four tones were suggested as possible candidates, including:

- ***special dial tone, positive indication tone, intrusion tone and congestion tone.***

Additional tones for other functions were also to be considered for standardisation if they were regarded as necessary, or others may be excluded. Technical and economic factors were also to be taken into account together with a limited evaluation for cost and feasibility and time scale for technical implementation of possible changes.

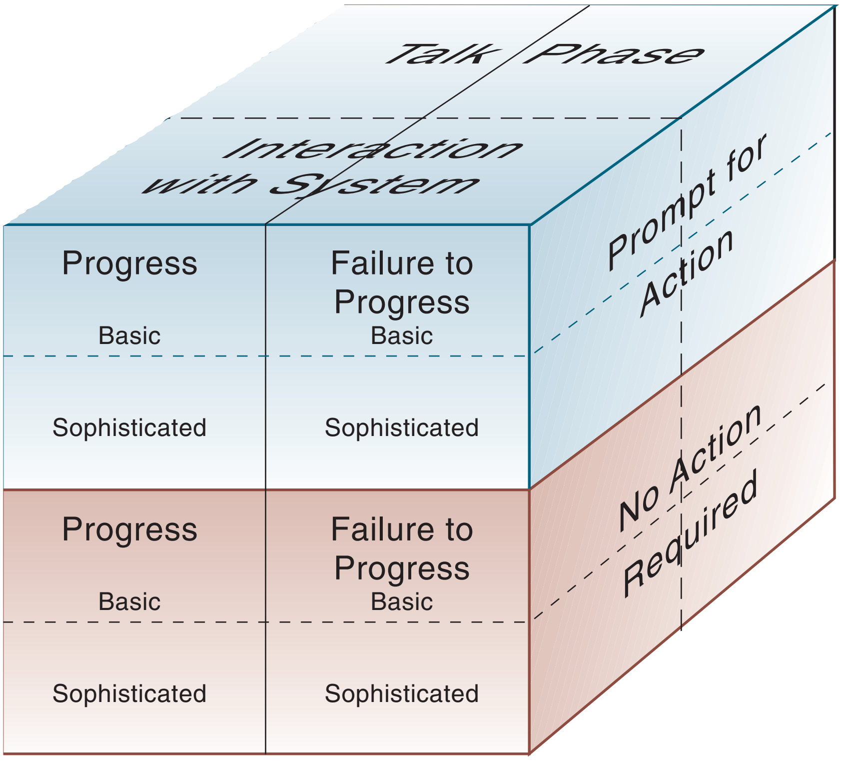


Figure 2 User goal based model

The Mandate, whilst giving broad guidance in terms of reviews that should be carried out, gave no particular hint as to how this mass of review material should be employed in any comparison process to evaluate the target tones in terms of possible need for harmonization.

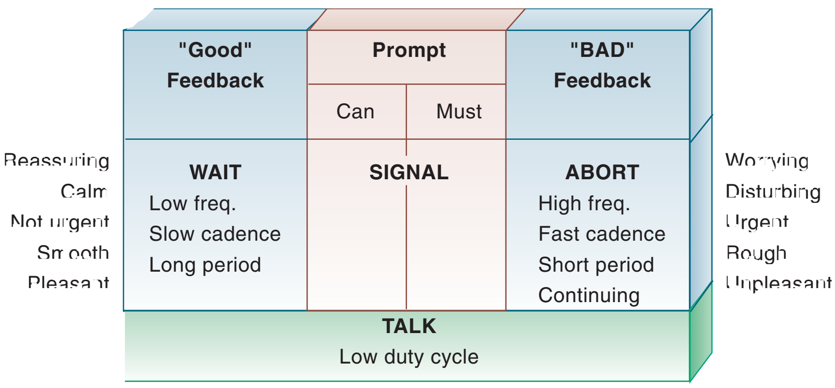


Figure 3 Tone characteristics model used in final audit

Ideally, Gagliardi's contentions should have been subjected to user testing to find out the degree of confusion or otherwise existing between tones and between user groups as he defined them (he defined two groups of users, those at home phoning to a different country and those abroad phoning home). This would have established a degree of usability which could have been compared with a normal human factors criterion of requiring 95 percent of the target population being able to discriminate accurately between the selected tones on 95 percent of occasions. The tests would have to include representative samples of the cultural mix in EU countries, age, sex and relevant telephone experience; a task of daunting proportions.

In human factors and practical terms, however, the only basis for criteria against which to test the tones were the data extracted from the reviews. During a brainstorming session by the study team many individual criteria were examined one at a time, such as the effect of tone characteristics like frequency and cadence, and the understanding of the meanings of tones. Attempts to combine this data in some way led to the concept of user modelling. From consideration of the sequence of tasks performed when making a telephone call, the first model was based on users' goals, and is illustrated in Figure 2.

However, this neglected the requirements of the telephone network, so that a task analysis was performed and consideration given to the indicate-control-indicate model already postulated for telephone user interaction. This gave rise to a series of models, which culminated in a final so-called human factors model, where not only were users alternative actions included, but also the relevant phases of a telephone call from the network point of view, as shown in Figure 3; for a better understanding of the content of each box, the model development is described later. When the human factors data about user perception of tones was also included, it seemed that a definitive model had emerged against which tones could be tested.

Residual doubts about the validity of the model remained, however, that it was not possible to satisfy without data from user testing, so that it was not appropriate to use the model alone as the basis for deciding on harmonization issues. Whilst it was not feasible to use the model as the sole criterion, it was possible to use the data from the reviews as well, in a linear comparison process termed an audit. In this way almost all items relevant to harmonization, not only human factors, would be given due consideration and form the basis for rational recommendations.

User perception and description of tones

A review of the human factors literature covered both general research and reference sources more specifically related to telecommunications. The findings were grouped into a number of areas as they relate to different aspects of the perception, discrimination and recognition of information tones. The review included some work on music and auditory icons, for instance, which allowed a broader view to be taken of harmonisation possibilities and options for change.

The review first established the basic facts about human perception of sounds. As being highly relevant in relation to tones in the telephone network, limits of human capacity for coding and tone discrimination were identified, as determinants of differentiation between tones and the limits to confusion. Considerations of recognition time and propensity to making errors were taken into account as being relevant to the efficiency of the telephone network, and work on warnings and urgency in similar context. The question of relating particular tones or their pitch to somewhat abstract concepts like *prompts* or *acknowledgement* was particularly relevant, but apparently a neglected area of research.

Two issues remained somewhat grey areas from the published data which are important in considering the validity of the model. One is the strength of association of certain tones with abstract concepts like *reassuring* or *disturbing* (see Figure 3 and later). This question is being studied currently at the Centre for Communication Interface Research (CCIR) in Edinburgh. CCIR are also performing other work with relevance to tones in exploring preferences for certain types of sound to represent categories of information, e.g. simple tones or rising or falling chimes. Another issue is the range of frequencies associated with qualitative expressions like high or low (frequency).

Successful telecommunications depends on the quality of the user/terminal/network interaction necessary for the purposes of establishing a call, modifying it (by invoking a supplementary service), or clearing it down. The quality of this interaction may be judged from the view of the network (or rather its designer) on the basis of whether the user responds to indications as desired, and from the view of the user on the basis of comprehensibility of such indications and the ease of use. A good system should meet the desires of both parties.

User-network interaction, task analysis and feedback

The procedural model in Figure 1 was based on a task analysis of the user-telephone network interaction, as a paradigm for user procedures. Each *indication* should consist of two components, an indication of the underlying *system status* and an indication of the *status of the required control*.

When this principle is integrated into a procedural task involving the system (terminal/network) progressing through a number of intermediary states in response to user control actions, the required indications become chained together. Any single indication given after a user's control action can thus have more than one function: it can be both feedback on the changes to the system/control status and a prompt for the next expected actions. This distinction is important when considering the roles that tones play within a user/terminal/network interaction, because the different roles may suggest differing characteristics for tones. For example, a crude distinction could be developed where *feedback tones* are slow and reassuring, but *prompt tones* are fast and alerting.

Other indications are not for feedback on a user's actions or to prompt a user to make a control input, but to prompt the user that they may have to change their behaviour to accommodate the changed/unexpected status of the current connection, e.g. Conference tone when a new party joins a conference.

In ordinary telephony, the number of options open to the user for controlling a call is limited and may be summarised as the SWAT model:

- **Signal** – to the terminal/network or service: includes going off-hook, dialling numbers, pressing function keys, etc.
- **Wait** – until the telephone, network, service or other user/s has done something
- **Abort** – return the terminal, network and/or service to idle
- **Talk** – all telecommunication including fax and data.

The feedback indications used to inform and prompt the user need to differentiate between which of these base level options is appropriate, and possibly also which is the most appropriate choice of the various signals for the user to send.

These complex interactions between users and networks were fully explored early in the project. They were more complex than first thought and led to a succession of models.

Table 1 Task Analysis of existing and proposed tones

Tone	Type	When	User Choices
Busy	Prompt	Set-up	Hang up or invoke SS - CCBS, UUS
Call waiting	Prompt	Talk	Ignore or Hang up to get new call or Attempt call control
Dial	Prompt	Set-up	Send or Dial address

Meaning of tones and feedback

Work in the area of audible signals suggests that specific characteristics of sounds impart specific meaning to the listener. For instance Pollack [5] has shown that tones with long cadences infer ‘goodness’ whereas tones with short cadences infer ‘badness’. In defining what attributes should be inherent in a set of telecommunications network tones it is necessary to consider what information the tones need to impart.

A task analysis of common user interactions with the telecommunications network suggested that the feedback given by tones needs to indicate success or failure in achieving the user’s desired goal; such tones either provide a prompt to stimulate an action from the user, or indicate that no further action is required. The information given by the tone needs to be sufficient to elicit the appropriate action from the user. Table 1 gives an example for three of the target tones.

Tones occurring at the end of a successful interaction can be considered to be feedback of the resultant network status, often requiring no action; tones occurring during a series of multiple interactions are usually prompts for action; in some cases tones have a dual purpose of both feedback and prompt. A further point to be considered is whether the tones are applied during the setting up of a call or service (when interworking with the network) or are applied later during the conversational phase of the call (talk or telecommunicate). Tones occurring during the conversational phase of a call are generally unexpected and indicate a change in the network status which may or may not be desired by the user.

Considering that only four actions in total are available to the user (Signal, Wait, Abort and Talk) and only three during the call set-up phase, it is tempting to consider whether more than three tones are required. This idea leads to the possibility of a taxonomy of tones into basic groups or families. Within any group of tones there can be simple, basic tones (reflecting the less informed user’s requirement for simple prompts) and sophisticated tones containing additional information permitting the knowledgeable user to adjust his actions to the prompt.

User goal based model

The multi-dimensional user goal based model in Figure 1 attempted to define the complete ‘set’ of network tones in terms of the following dimensions:

Dimension One – Action Required / Not Required: On hearing the tone in question the user is prompted either to carry out a specific action in order to move closer to achieving their goal (which could be setting up a conversation, setting up a network service, etc.) or to remain passive, awaiting the next phase in the process.

Dimension Two – Progress / Failure to progress: On hearing the tone in question, the user is aware whether or not they have moved closer to achieving their goal.

Dimension Three – Interaction with the system / Talk phase: Tones occurring in the interaction phase are normally expected feedback tones. Tones occurring during the Talk phase are generally unexpected.

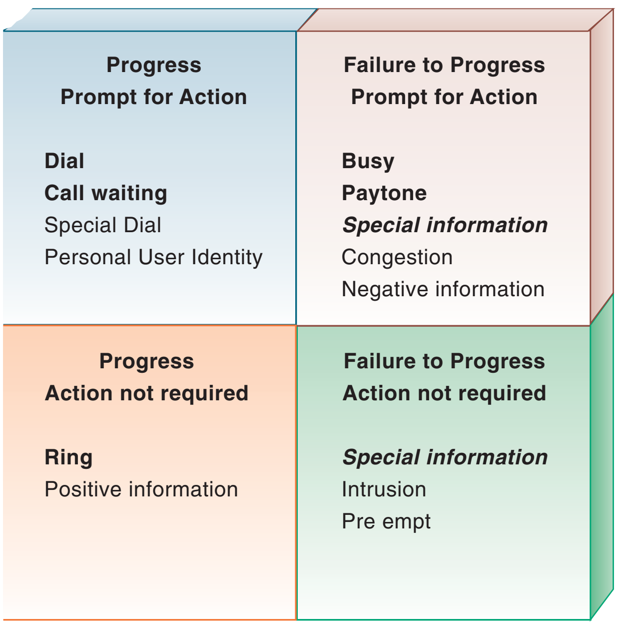


Figure 4 Tone assignments for dimensions 1 and 2

Dimension Four – Basic tone / sophisticated tone: Dial tone can be considered to be a basic tone, giving an indication of reaching the desired system status and also prompting for a number input. Special dial tone would be classified as a sophisticated tone giving additional information on the system status to an expert user.

The benefits of such a model/taxonomy are as follows:

1 If it is possible to associate each of the dimensions with a unique tone characteristic, then the specification of a new tone (or tone set) will be facilitated.

2 Categorisation as above indicates groupings and/or individual tones which may be considered in more detail for combination with others, or deletion.

3 If the tone *characteristics* define whether action is required and whether progress is being made towards the goal, then in the worst case situation the user may still achieve their goal without recognising ANY INDIVIDUAL tones, only their characteristics.

Categorisation and characteristics of tones

The next step was to consider only the first two dimensions of the model and assign to them the tones defined in the Mandate (and some others), with the tones shown in **bold** being those defined as priority.

Figure 4 shows the assignments which were reasonably straightforward, and gave some confidence in the applicability of the model. The model gave somewhat indeterminate results for only one tone. Special information tone as defined by ITU-T [6] is used to indicate call failure. In some cases it stands on its own and requires that the caller acts to release the call. In other cases it is followed by a message and requires the caller to wait for the message before deciding on a course of action. It is thus context dependent and is *italicised* to show that it does not fall firmly into a single classification.

When attempting to apply the model it was necessary to postulate tone characteristics appropriate to the various dimensions of the model. The following were identified as possibilities for mapping onto the tone model defined above:

Frequency: High/low, single/multiple, simultaneous/sequential, rising/falling, discrete steps/ continuous.

Complexity: Harmonic content.

Cadence: Fast/slow, Duration, Duty Cycle, Continuing, Once/ twice.

Intensity: Comparative.

Attempting to define which dimensions to apply to the "progress" axis and which to apply to the "action" axis gave rise to the following interim proposals:

In the axis 'progress/failure to progress';

Progress:

- An indication of satisfactory progress should be a 'good' tone with reassuring characteristics having a calming effect.
- Low frequency, harmonious, slow cadence, Pollack [5], Pollack and Ficks [7], Heberle [8], rising frequency?

Failure to progress:

- An indication of failure or imminent failure to progress should be a 'bad' tone with disturbing characteristics having a stimulating effect.
- High frequency, discordant, rapid cadence [5,7,8].

In the axis 'Action required/action not required';

Action required

- A tone indicating action required should awaken the user, convey urgency and be persistent so as to invite action
- High frequency, discordant, rapid cadence, Patterson & Milroy [9], continuing (until action performed)?

Action not required

- A tone indicating action not required should lull the user into inaction
- Low frequency, harmonious, slow, Patterson & Milroy [9], cadence?

Characteristics for these tones from the literature were then placed into the four cells relevant to the first two dimensions of the model (the tones not allocated from experimental evidence being *italicised*), shown in Table 2. This then shows that two cells are more or less concordant but two are also contradictory. This could suggest that this level of coding discrimination was too simplistic, that the context in which the published information was derived was not relevant, or it could illustrate a conflict between the needs of the system designer and those of the user.

Table 2 Assignment of tone characteristics from the literature

	Progress	**Failure to progress**
Action required	Progress – choose from: Low frequency, *Harmonious,* Slow cadence, *Rising frequency*	Failure to progress – choose from: **High frequency**, ***Discordant*** **Rapid cadence**, *Falling frequency*
	Action required – choose from High Frequency, *Discordant,* Rapid cadence, *Continuing*	Action required – choose from **High Frequency**, ***Discordant***, **Rapid cadence**, *Continuing*
Action not required	Progress – choose from: **Low frequency**, ***Harmonious,*** **Slow cadence**, *Rising Frequency*	Failure to progress – choose from: *High frequency, Discordant,* Rapid cadence, *Falling frequency*
	Action not required – choose from **Low Frequency**, ***Harmonious,*** **Slow cadence**	Action not required – choose from Low Frequency, *Harmonious,* Slow cadence

A number of alternative models were subsequently investigated, and a number of unsuccessful attempts were made to select a subset of tone characteristics that would satisfy the model in all aspects, but it was finally concluded that as the network was not itself user goal related, the model in all its variants investigated did not fully allow mapping of all possible goals, or adequately take into account the limitations on the prompts and information to be conveyed to the user.

Another model was designed to allow classification of tones by system state, their characteristics, and presentation to the user. It used the original concept of OK / Not OK, gave a general indication to the user of call progress towards an (assumed) goal, indicated the phase of the call with the addition of talk/non-talk and made a distinction between Feedback, Prompt 'Can' and Prompt 'Must'. (Feedback is the simple provision of information on system status, Prompt 'Can' suggests some choice for the user and Prompt 'Must' demands action from the user in order to prevent failure of the call.) This version was further simplified, rearranged and populated with tone characteristics derived from the human factors literature, and the subjective descriptions associated with them for each phase according to the SWAT model, and this led to the model used during audit, shown in Figure 3.

The final model encompasses virtually all of the dimensions analysed as characteristics of the user-telephone interaction. Some assumptions relate to user goals in this interaction, and others to system goals, because a single model cannot represent them all. Tones assigned to the model appear to be in their right places. If the characteristics of known tones as assigned to the model in Figure 2 are examined, they are found largely to conform with the literature-derived figures in most cases. One exception is Dial Tone, whose functional definition does not make it completely clear whether it should be considered as a prompt or a feedback tone. Many users merely accept dial tone as feedback indicating that the network (or terminal) is working, or it may also be interpreted to suggest that the tone should be within the family 'Prompt Can'.

Table 3 *Possible values for tone characteristics*

	Frequency	Rhythm	Period	Other
'Good'	Low < 500 Hz	Slow – 1/s	Long – 3 s	
'Bad'	High > 500 Hz	Fast – 2/s	Short – 1 s	Continuing
	Non-intrusive	**Talk Phase**	Intrusive	
	Duty cycle < 5 %		Duty cycle > 15 %	

Nevertheless, there is some confidence in using the model to assess the conformance of each tone to the model, and to use it to identify the characteristics of a new tone for a given function in future, provided its functional description can be matched to a cell in the model, as discussed in the audit later.

User requirements

In addition to the use of modelling to identify potential attributes of tones relevant to the type of feedback required, the reviews carried out early in the study identified some significant user requirements, based on specific human user characteristic, such as the auditory sense, or memory ability for procedures and meanings of information tones, or general requirements for the usability of the network at a national or international level. Some requirements were identified in the first place by Gagliardi, others during the derivation of the model, and others from the literature or from good human factors practice. Some of these user requirements were used as general criteria during the comparison process, or audit, described below.

Use of the model

The final model encompasses virtually all of the dimensions analysed as characteristics of the user-telephone interaction. Some assumptions relate to user goals in this interaction, and others to system goals, because a single model cannot represent them all. Tones assigned to the model appear to be in their right places.

There is therefore some confidence in using the model to assess the conformance of each tone to the model, and to use it to identify the characteristics of a new tone for a given function in future, provided its functional description can be matched to a cell in the model. This matching is part of the audit procedure developed and described to examine each of Gagliardi's six priority tones, in order to sensibly determine how they might be 'harmonised' in the future.

For the purposes of this audit, values needed to be assigned for the technical characteristics described in the model. Table 3 gives a preliminary assessment of these values (it is hoped that they can be conformed at some later stage by suitable experiment).

Consideration of Gagliardi tones for harmonisation

The final stage of the team's work was to assess each target tone described in the mandate in order to recommend possible ways of harmonising them in the EU environment. Data available at this stage included results of the reviews and the analysis of European tones, as well as potential for guidance from the final models. This assessment was carried out by means of an audit of each tone, following a detailed procedure consisting of ten steps:

1 The target tone is first named and defined by its functional description, i.e. what it is intended to convey to the user about a call. At this stage it is possible to identify some tones where it is not clear either what their function is, or that they are using the wrong description or name.

2 The next step is to assign the tone to a position in the HF model, based on its functional description.

3 Step 3 then defines the characteristics of the tone postulated by the HF model.

4 Step 4 identifies the range of characteristics reported for the target tone by the chosen group, i.e. Europe. These data may be inspected for coherence, spread of certain characteristics, and any significant modal value.

5 Step 5 identifies a modal set of characteristics for the tone. By inspection, it may be possible to identify a 'mode' group within any one tone – that is, the tone characteristics which the largest number of countries report using. Such a mode may be extracted from any set of tones grouped by economic, regional or geographical boundaries.

6 Step 6 inspects the existing standards for that tone, both for network and terminal generation.

7 The relevant human factors issues on the users' expectations, the potential for confusion and any usability issues are reviewed next. The tone characteristics are compared for possible confusability with other tones in the same group or any other group, and usability in relation to the task the user is carrying out. For the purposes of the audit, only an expert estimation can be made of usability in terms of likely confusion with other examples in the tone group for users in Gagliardi's groups A and B (those visitors to a foreign country dialling internally and users dialling into another country). Doubtful cases can only be settled by usability testing.

8 Five options for harmonisation or change, labelled a–e, are now examined in some detail:

Option a: Use the HF model. In this case harmonisation could be achieved by countries aligning the characteristics of their tone with those proposed by the model. In the case of a gross misfit, the audit would try to determine whether any change could be justified in terms of improvement in usability versus cost.

Option b: Use the existing standards. In this case, harmonisation could be achieved by countries changing some aspects of a tone's technical characteristics to align it with a specific international standard.

Option c: Use the European mode. As explained above, harmonisation could be achieved by countries changing harmonizing their tone's characteristics to those most commonly occurring in Europe (or in any other grouping).

Option d: Do something different. In an exceptional or new case, it may be desirable to create a tone with characteristics completely outside an existing range, in order to make it truly distinguishable, or a characteristic 'earcon' may be created.

Option e: Do nothing. This would imply that the tone was usable, would not confuse users in Gagliardi's Categories A or B, and was conformant with known definitions and standards. In other words, the tone was in effect already harmonised.

No other Options are considered, there are other options for change which could have been examined, for example:

- Different calculations of the European mode – e.g. by weighting countries by reference to their number of subscribers, or their volume of traffic, or by their ETSI voting rights.
- The consideration of a World Mode, however calculated.
- The specific use of new tones or earcons.

However, these five (a–e above) represent the key choices that may be made.

9 The penultimate step reviews the implications for each option (a–e), including crude estimates of cost and benefits and the ease with which a possible migration path may be established as part of an implementation plan.

10 The final step derives a set of recommendations in priority order with respect to the harmonisation of each tone.

An example of one such audit, of *ringing* tone is given in Annex 1.

Conclusions

There were two satisfactory outcomes from the study. One was the development of a model, based mainly on empirical findings and a task analysis of the users' activities in making a telephone call. Applying some research evidence and testing known tones to the model appears to confirm its validity, but some further testing should be done, particularly with respect to any possible predictive value the model may have.

Secondly, the many detailed analyses done, and the data extracted from the human factors literature about human perception of tones, lay the foundation for an audit process as a suitable logical basis for recommendations for harmonization of network tones. In particular, the systematic consideration of human factors issues, such as usability assessment, would ensure that if the processes were to be used for new tones, they would not only fit the model, but be acceptable and useful to representative users.

The analysis of the tone data showed the strong dependence at present on using different simple cadences to differentiate between tones, while neglecting greater use of frequency. At the same time a wide interpretation of the functional definitions of some tones has lead to the sometimes inappropriate applications in networks, to a degree likely to cause confusion. The study thus confirmed partly the concern by Gagliardi with possible confusion and error when using the telephone within the EU countries. The recommendations, containing mostly minor changes to achieve harmonization are set out in full in the ETSI Technical Report.

The recommendations particularly emphasize the need to harmonize the functional definition of some tones to ensure their correct application and understanding by users and for the cessation of incorrect usage. It is seen as especially important that the same or similar tones are not used within a network to signify different meanings or services.

In the case of *dial* tone, it was proposed that a consistent use of a continuous tone of single or dual frequencies be adopted, and gradual change to a harmonized continuous tone of 425 Hz.

For *ringing* tone proposals included harmonization of cadence and frequency, and that requirements for a number of possible special ringing tones be developed, such as for special services, videotelephone call, voice mail, etc.

At the same time, options should be provided for new dial and ringing tones, based on the model, to assist the provision of corporate identity in the tone.

It was suggested that benefits might come from introducing a new harmonized high frequency *busy* tone in preference to existing usage. Similarly, for *special information tone*, a new tone might be developed in line with the HF model in the call failure group to indicate *number unobtainable* or *invalid service*.

It must be emphasized that the recommendations from ETSI to the European Commission remain just that at the time of publication, and no action has been proposed or discussed.

Acknowledgement

The authors are grateful to the Director-General of ETSI for permission to publish this paper which is based on work performed on behalf of ETSI under CEC Mandate BC-T-308.

References

1 Karlin, J. *Summary of results from laboratory studies of audible tone signals, including data from Germany, Japan, The Netherlands, United Kingdom, and United States.* CCITT WP XIII/1, Temporary Document No. 6-E. Tokyo, July 1967.

2 Gagliardi, D. *Report on the audible tones in the telephone service in the EC countries.* Report from contract 48159, DGXIII, European Commission. Brussels, April 1993.

3 ETSI TR 101 041-1 ed. 1. *Human Factors (HF) : European harmonization of network generated tones : Part 1 : A review and recommendations,* Sophia Antipolis, 1997.

4 ETSI TR 101 041-2 ed. 1. *Human Factors (HF) : European harmonization of network generated tones : Part 2 : Listing and analysis of Euro-*

pean, World and Standardized tones. Sophia Antipolis, 1997.

5 Pollack, I. The information of elementary auditory displays. *Journal of the Acoustical Society of America,* 24, November 1952.

6 ITU-T Recommendation E.180. *Technical characteristics of tones for the telephone service.* Geneva, 1988.

7 Pollack, I, Ficks, L. The information of elementary multi-dimensional auditory displays. *Journal of the Acoustical Society of America,* 26, 1954, 155–158.

8 Heberle, W. Human aspects of audible tones for controlling telephone user's behaviour. *Proceedings of 6th International Symposium on Human Factors in Telecommunications, Stockholm,* 1972.

9 Patterson, R D, Milroy, R. *Auditory warnings on civil aircraft : The learning and retention of warnings.* Cambridge, UK Medical Research Council, Applied Psychology Unit, 1980.

Donald M. Anderson is currently working as Research Ergonomist with Rikshospitalet (National Hospital), Oslo, and as the Principal Consultant, Man Machine Technology AS. He has 45 years experience in ergonomics and his interests include standards, maritime ergonomics, product design and development and aids for the handicapped.

e-mail:
mantech@login.eunet.no

Annex 1 Examples from analysis of European tones

1 Dial tone

Functional definition

The auditory indication to be presented to a user to indicate that a network connection is available and ready to receive call information and inviting the user to start sending call or service related information. (This definition is consistent with ITU-T Recommendation E.182.)

Table of European Dial Tones

	Cadence	Period	Frequency	Country
	None reported			Malta
1	**0,33** – 0,33 – **0,66** – 0,66	1,98	425	Czech Republic, Slovakia
2	**0,2** – 0,2 – **0,6** – 1,0	2	425	Italy
3	**0,7** – 0,8 – **0,2** – 0,3	2	425	Slovenia, Yugoslavia
4	**0,2** – 0,3 – **0,7** – 0,8	2	425	Croatia
5	**0,2** – 0,3 – **0,7** – 0,8	2	425 or 450	Greece 1 (ITU)
6	***0,2*** *– 0,8 –* ***0,7*** *– 0,3*	2	*425 or 450*	Greece 2 *(300 085)*
7	**0,25** – 0,3 – **0,7** – 0,8	2,05	425	Bulgaria 1
8	**0,25** – 0,75 – **0,75** – 1,0	2,75	425	Bulgaria 2
9	Continuous		350+400	Gibraltar
10	Continuous		350+440	United Kingdom
11	Continuous		350+450	CI Jersey, Cyprus
12	Continuous		400 or 425	Portugal 1 (ITU)
13	Continuous		*400*	Portugal 2 *(Gagliardi)*
14	Continuous		400 or 425 or 450	Ireland 1 (ITU)
15	Continuous		*400 or 450*	Ireland 2 (Gagliardi)
16	Continuous		400 or 450	Romania
17	Continuous		420 or 450	Austria
18	Continuous		425	Albania, Denmark, Estonia, Faroe Islands, Finland, Hungary, Iceland, Lithuania, Poland, Portugal 3 (300 085), Russia, Spain, Sweden, Switzerland, Norway
19	Continuous		425 or 450	Belgium, Germany, Luxembourg, Netherlands
20	Continuous		440	France
21	Continuous		450	Turkey

2 Ringing tones

Functional definition

The auditory indication to be presented to a user to indicate that a connection has been made and that an alerting signal is being applied to the called terminal or service. (This definition is consistent with ITU-T Recommendation E.182.)

Table of European Ringing Tones

	Cadence	Period	Frequency	Country
1	**0,4** – 0,2 – **0,4** – 2,0	3	400 or 450 or 425 x 25	Ireland 1
2	**0,4** – 0,2 – **0,4** – 2,0	3	400 x 25	Cyprus 2
3	**0,4** – 0,2 – **0,4** – 2,0	3	400+450	Gibraltar, CI Jersey
4	**0,4** – 0,2 – **0,4** – 2,0	3	400+450 or 450 x 25 or 425 x 1762/3	United Kingdom, Malta
5	***0,4** – 0,2 – **0,4** – 2,0*	*3*	*425*	Ireland 2 (Gagliardi)
6	**0,8** – 3,2	4	425	Lithuania, Russia 1
7	**1,0** – 3,0	4	400 or 450	Norway 2
8	**1,0** – 3,0	4	425	Cyprus 1
9	**1,0** – 3,0	4	425 or 450	Belgium
10	**1,5** – 3,0	4,5	425	Spain
11	**1,83** – 3,0	4,83	425	Bulgaria 3
12	**1,2** – 3,7	4,9	425	Hungary
13	**1,0** – 4,0	5	425	Albania, Bulgaria 2, Croatia, Czech Republic, Denmark, Estonia, Faroe Islands, Finland, Italy, Luxembourg 1, Norway 1, Poland, Russia 2, Slovakia, Slovenia, Yugoslavia 2
14	***1,0** – 4,0*	*5*	*425 or 450*	Germany 5 (Gagliardi), Greece, Netherlands
15	**1,0** – 4,0	5	425 or 500	Switzerland
16	***1,0** – 4,0*	*5*	*450*	Luxembourg 2 (Gagliardi)
17	**1,5** – 3,5	5	440	France
18	**1,2** – 4,7	5,9	425	Iceland
19	**1,0** – 5,0	6	400 or 425	Portugal
20	**1,0** – 5,0	6	420 or 450	Austria
21	**1,0** – 5,0	6	425	Sweden 1
22	**1,0** – 5,0	6	425 or 450	Germany 3
23	**2,0** – 4,0	6	400 x 16 or 450 x 25	Romania
24	**2,0** – 4,0	6	450	Turkey
25	**0,25** – 4,0 – **1,0** – 4,0	9,25	425 or 450	Germany 1
26	**0,5** – 4,0 – **1,0** – 4,0	9,5	425 or 450	Germany 2
27	**1,0** – 9,0	10	400	Sweden 2
28	**1,0** – 9,0	10	425	Bulgaria 1
29	**1,0** – 9,0	10	450	Germany 4
30	**1,0** – 9,0	10	450 x 25	Yugoslavia 1

Annex 2 Examples of tone audits

Audit of tone 'DIALLING'

1 Dial tone – Functional definition

The auditory indication to be presented to a user to indicate that a network connection is available and ready to receive call information and inviting the user to start sending call or service related information. (This definition is consistent with ITU-T Recommendation E.182.)

2 Position in HF model for dial tone

This functional definition places the dial tone in the HF model within the family of Feedback + System State OK and at the "Good" end of our characteristics continuum. It may also be interpreted to suggest that the tone should be within the family Prompt Can.

3 HF model characteristics for dial tone

- Low frequency < 500 Hz
- Slow rhythm < 1/s
- Long period > 3 s

4 Current European practice for dial tone

Table of dial tones used in Europe

Cadence	Frequency	No. of Countries
Continuous	350 + 400	1
Continuous	350 + 440	1
Continuous	350 + 450	2
Continuous	400 or 425	1
Continuous	400 or 425 or 450	1
Continuous	400 or 450	1
Continuous	420 or 450	1
Continuous	425	14
Continuous	425 or 450	4
Continuous	440	1
Continuous	450	1
0,165 – 0,165 – **0,66** – 0,66	425	a
0,3 – 0,33 – **0,66** – 0,66	425	1 + a
***0,2** – 0,2 – **0,6** – 1,0*	*425*	1
0,7 – 0,8 – **0,2** – 0,3	425	2
0,2 – 0,3 – **0,7** – 0,8	425	1
0,2 – 0,3 – **0,7** – 0,8	425 or 450	b
0,2 – 0,8 – **0,7** – 0,3	425 or 450	b
***0,25** – 0,3 – **0,7** – 0,8*	*425*	c
0,25 – 0,75 – **0,75** – 1,0	425	c

Note: Letters are used for countries which report two or more different tone characteristics for the same function. See Part 2 for the fuller details.

5 European mode for dial tone – Characteristics

Continuous cadence, 425 Hz Single frequency.

6 International standards for dial tone

Table of International standards for network generated dial tones

Network generated tones			
Source	**Cadence**	**Fre-quency**	**Preference for any update of equip-ment**
ITU-T Rec. E.180	Continuous	Single f, range 400 – 450 Com-bined, up to 3f (1f range 340 – 425, 2f range 400 – 450)	Continuous cadence at single frequency of 425 ± 15 Hz
CEPT T/SF 23	Continuous	400 – 450	
CEPT T/CS 20-15	Continuous	425 ± 15	

Table of International standards for terminal generated dial tones

Terminal generated tones		
Source	**Cadence**	**Frequency**
prETS 300 245-7 (ISDN Terminals)	Continuous or National Tones	425 or National Tones
ETR 187 (General)	Continuous	425
GSM 02.40 (Dial tone is not normally required, but if it is provided)	N/A Continuous	N/A 425 ± 15
ETR 294 (TETRA – no characteristics defined)	N/A	N/A

7 Human factors issues for dial tone

Expectations

A continuous low frequency tone of single or dual frequencies

Confusability

HF expertise would suggest that in the absence of the facility to compare two frequencies

- few users (may be as low as 10 %) would report a difference between two continuous single frequency tones, when the variation of tones used is within ±25 Hz;
- more users (may be as many as 50 %) would report a difference between continuous multi-frequency tones and continuous single frequency tones;
- most users (may be as high as 95 %) would report a difference between discontinuous single or multi-frequency tones and continuous single or multi-frequency tones.

Usability

In practice most users of unfamiliar networks may assume that any auditory signal presented on requesting connection to a network can be taken to be the invitation to send call or service related information.

The alternative – Silence – would be taken to indicate no connection is available.

8 Review of options for dial tone

a) Use the HF Model

If the dial tone is correctly located in the 'Feedback/System OK' family, the characteristics should be located towards the lower frequency slower end of the range. If it should be located in the 'Prompt can' family, it should lie more towards the higher frequency fast end of the same characteristics.

In this context, if continuous is taken as the extreme of slow cadence, the HF model requirements are similar to the requirements of the standards and also the existing European mode.

b) Use the standards

The characteristics defined in the current ITU-T, CEPT and ETSI standards are all coherent with each other. They all define a Continuous tone with differing tolerances on the acceptable frequencies. There are two EU and a number of other European countries that do not comply with the loosest standard ITU-T, these are identified in subclause 8.2.1.9. The key change required for compliance with ITU-T is for countries with Discontinuous Tones to migrate to a Continuous Tone. The changes required for compliance with the tightest standard are the same as for converting to the European mode.

c) Use the European mode

The characteristics of the modal dial tone provide a good fit to the HF model, but there are a number of EU and other European Countries with Dial Tones that do not precisely match these characteristics; these are identified in subclause 8.2.1.9. The key changes required are for some countries to migrate from a Discontinuous to a Continuous tone, some to migrate from a dual to a single frequency and for others to change the frequency.

However, the HF evidence suggests that:

- the differences within the range of single frequencies used is unlikely to cause any confusion;
- the differences between the dual frequency and single frequency tones is unlikely to cause much confusion;
- the use of a discontinuous tone with a short period is the most likely source of confusion to Category A users.

d) Do Something different

There is no human factors or technical reason for such a change to the public network service. Where there may be a requirement to express a 'corporate identity' through a 'unique' dial tone, any proposed tone or earcon would need to be tested for usability.

e) Do nothing

The spread of characteristics of existing European dial tones is not large, and there are few confusion or usability issues. Where these may be expected i.e. in the case of discontinuous tones and perhaps also in the case of dual frequency tones, the impact will be felt most by Category A users (travellers temporarily using another network).

9 Implications of options for dial tone

Option a – Use the HF model

No changes required.

Option b – Use the Standards

Change to a Continuous Cadence.

EU

Greece (change cadence to continuous, and maybe change NU tone for Paging)
Italy (change cadence to continuous, and change Special Dial tone)

ETSI area

Bulgaria, Croatia, Slovakia, Slovenia & Yugoslavia (change cadence to continuous)
Czech Rep. (maybe change cadence to continuous)

Option c – Use the Mode

Change to a Continuous Cadence and a Single Frequency (Based on Continuous at 425 Hz).

EU

Greece (change cadence to continuous and maybe change f on some exchanges 450 to 425 and maybe change NU tone for Paging)
Italy (change cadence to continuous; and change Special Dial tone)
United Kingdom (change f from 350 + 440 to 425)
Insignificant changes would be required in Austria, Belgium, Germany, Luxembourg, Netherlands, France, Ireland, Portugal

ETSI area

Bulgaria, Croatia, Slovakia, Slovenia, & Yugoslavia (change cadence to continuous)
CI Jersey & Cyprus (change f from 350 + 450 to 425)
Czech Rep. (maybe change the cadence on some exchanges to continuous)
Gibraltar (change f from 350 + 400 to 425)
Insignificant changes would be required in Rumania & Turkey

Option d – Do something different

The implications of this course of action are dependent on just how different the 'different' is. At least it would imply extensive testing of the proposed tone or earcon.

Option e – Do nothing

There are no additional costs in this course of action but benefits will be foregone, particularly to terminal equipment manufacturers and users.

10 Recommendations for dial tone

1 To encourage the rapid harmonisation of the functional use for dial tone and to encourage the cessation of the use of a similar tone to indicate other system functions, e.g. positive indication.

2 To encourage the rapid change to the consistent use of a Continuous dial tone of single or dual frequencies.

3 To encourage the gradual change to a single harmonised dial tone based on the Continuous tone of 425 Hz, when planned changes to the public networks make it possible.

4 To encourage National Regulatory Authorities (NRAs) to require new public network operators providing services which require the provision of a dial tone to comply with either their existing national dial tone characteristics or to provide a dial tone with a Continuous tone of 425 Hz.

5 To permit new dial tones based on the HF model that may assist the provision of a 'Corporate Identity' to Public Network Operators or Private Corporate Networks. NRAs should ensure that any new dial tones are properly regulated, are usable, are compliant with the HF model and are significantly distinguishable from network tones having other meanings.

The ergonomics of teleworking

JOHN W. BAKKE

> *An inspection of the concrete operation of any cooperative system shows [...] that the physical environment is an inseparable part of it.* [1]

A characteristic feature of teleworking is the geographical dispersion of an organization. This dispersion has profound and distinct implications for the practice of teleworking, compared to other work arrangements. One set of implications relates to the organization of work, like co-operation, co-ordination and management. Another set of implications relates to what might be called the ergonomics of teleworking, the working environment challenges of dispersed work.

1 Do you take your body to work?

The question "Why bring people to work, when they can work where they are?" may serve as a motto for the field of teleworking. Originally, the idea of teleworking was motivated by a concern for the oil-dependency of western economies, and the potentials for reducing road-based transportation [2]. Later, the theme of teleworking was incorporated into the literature on office automation, and into visions of the development of more humane and democratic future societies [3, 4, 5].

Since the inception of the idea, a large body of literature on teleworking has been produced, where a few predominant themes have dominated the discourse: In addition to a plethora of descriptive accounts in the popular media, one finds descriptions of technological developments, potentials and solutions, a large number of teleworking guides and handbooks, and several case-studies, usually pilot studies, in the field of organization theory.

A decisive background for the idea of teleworking is the development of information- and communication technologies, allowing, in principle, a large number of work tasks to be performed at an arbitrary location. 'Information work' is considered especially appropriate for telework, since it is easier to move information (or 'bits') than bodies or products (or 'atoms' – see [6]). The interest in teleworking is further motivated by intensified competition and a rapidly changing competitive environment, new organizational forms and management techniques, and the increasing pressure on the family's and the individual's time schedules [7].

In this presentation, selected themes in the ergonomics of teleworking will be explored. Teleworking is potentially an interesting arena for studies of ergonomics, motivated by the specific organizational characteristics of this way of organizing work, as well as the sheer number of teleworkers. Although the concept of 'teleworking' covers a variety of work environments and work practices, it will be argued that there are certain, well defined, ergonomic challenges of teleworking.

In the first sections, a brief presentation of teleworking, and glimpses from ergonomics, organization theory, architecture and design will be given. This provides a background for a discussion of the ergonomical challenges of teleworking, through an exposition of certain characteristics of this way of organizing work. In this section, it will also be argued that studies of the ergonomics of (tele)working may serve as a link to more general themes in studies of organizations, like gender, culture, and the processes of modernization. In a final section of the essay, implications for further research will be outlined.

Figure 1 A number of work tasks may be performed equally well at a centralized workplace ...

2 Situating telework

'Teleworking' is not a well defined concept. While there are several examples, which doubtless do count as instances of teleworking, there is a large number of work situations that may – or may not – be counted as teleworking. For the purposes of this essay, an encompassing concept of teleworking will be deployed – mainly understood as *work at a different location than the employer's*. This definition is considered beneficial, by allowing a wide variety of practices within a field of rapid development to be included in the analysis. A more precise definition may be obtained by specifying threshold levels for regularity and volume, technology deployed, and the existence of an agreement between employer and employee, while bearing in mind that such threshold levels are arbitrarily given [8, 9].

2.1 A brief history of telework

For a couple of decades, teleworking has been a theme in visions of the future of work. In a now classic study of the "telecommunication-transportation tradeoff" [2], *telecommuting* was proposed as one solution for the oil-dependence of the western economies. Aided by communications technologies, the argument went, one might reduce road-based commuting by working one or more days per week at home, or in a neighbourhood telework centre, thereby reducing travel and traffic congestion, in particular in urban areas. Later, the term *teleworking* was introduced as a general term for "ANY form of substitution of information technologies [...] for work-related travel" [7]. The concept telecommuting has, however, remained in use as the general term, at least in the American context, while in Europe, the concept teleworking has become accepted as the general term [8].

As indicated above, the theme of teleworking was soon incorporated in more elaborate visions of the future of work, and the future of society at large. The motivations for establishing teleworking schemes went beyond the 'substitution for transportation'-argument, and teleworking was presented as a way of integrating work and family life [4], a way of relocating work to rural areas [10], a way of offering peace and quiet for work groups [11], a way of capitalizing on investments in information technologies [12], and a way of slimming expensive office space [13].

In the study of the telecommunications-transportation trade-off [2], teleworking centres, or 'business centres', was given a prominent place, whereas home-based teleworking got less attention – although the authors referred to the media awareness of home-based work. It was the futurist Alvin Toffler who most forcefully formulated the ideal of home-based teleworking, in what he called the *electronic cottage*. The result, he proposed, would be a 'home-centred society'. In later studies, home-based teleworking has come to the forefront, alone, or in combinations with office-based work and mobile teleworking (see [7] and [14], among other sources).

From its early beginnings, the discourse on teleworking has had an international orientation. The initial studies were made in the United States, but very soon the idea spread to other countries as well, while deploying the US, and later the British and Swedish experiences as inspirations [8]. Now, information policies like the *Bangemann Report* promote teleworking, and the goal has been formulated to reach two million teleworkers in the European Union by the turn of the century [12].

While being internationally inspired, one also finds national characteristic traits. For instance, local teleworking centres, or 'telecottages' have been proclaimed as a Scandinavian solution, retaining the positive aspects of teleworking, like short commuting travels and presence in the local community, while avoiding perils like isolation or an unfavourable bargaining position relative to the employer [15].

To a large extent, the literature on teleworking has a pro-teleworking bias, often in mixtures of descriptions and prescriptions. The majority of studies are simple assertions of the possibility of teleworking, and interviews with happy and satisfied teleworkers and teleworkers' managers and families. In addition, one finds a large number of case-studies from single or a few companies. There are few comprehensive and even fewer comparative studies of this way of organizing work (for a discussion of the literature on teleworking, and its often dubious quality, see [8]).

The motivations for establishing teleworking schemes vary considerably. It seems reasonable to assume that for teleworking schemes to become successful, they have to some extent to benefit the parties involved – most directly, the employer and the employees in question. When discussing motivations, one may distinguish between work-related arguments, and non-work-related arguments for teleworking. Further, one may assume that employers are more interested in the work-related implications of teleworking, like increased productivity, whereas employees are interested in both work-related aspects like peace and quiet for concentration away from a hectic office, and non-work-related aspects, related to commuting, or the management of the boundary between work and non-work.

Certain aspects of teleworking may be conceptualized *a priori*, like the necessity of overcoming the distance between the teleworker at the remote workplace, and his or her colleagues, managers and other relevant parties [7], or the challenge of managing the boundary between home and work, when working at home [16]. With teleworking as such a wide-ranging phenomenon, there are, however, few 'inherent' characteristics of this way of organizing work. For instance, some managerial strategies like 'management by walking around' seems to be inappropriate in teleworking situations. This will, however, only be the case if the teleworker spends the majority of his

Figure 2 ... or at a home-based office (Source: Information brochure from Telenor)

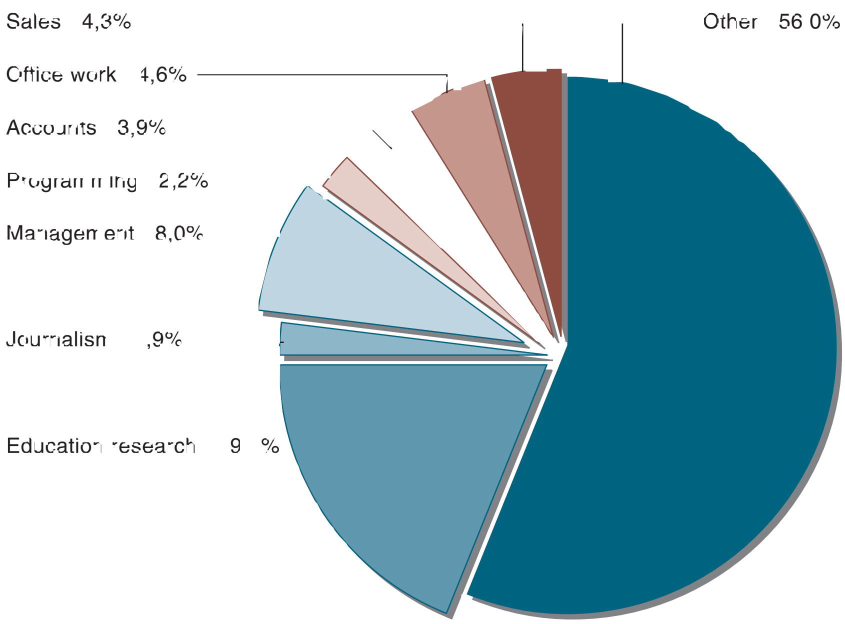


Figure 3 Primary occupation of home-based teleworkers

or her working time at a remote office, not if teleworking takes place intermittently, or after ordinary working hours.

2.2 Teleworking in Norway

In Norway, teleworking was put on the political agenda in the early 1980s, in an Official Norwegian Report on the future of telematics [15]. In this report, teleworking was primarily discussed as an instrument for decentralization and job-creation, and knowledge transfer to rural areas. While stating that home-based teleworking already was the predominant form, the report argued that this form of teleworking would not provide a satisfactory work environment, according to existing laws and regulations. Instead, telework centres are presented as a preferable solution, providing a better psycho-social work environment. A number of teleworking centres were established, mainly on initiative of the authorities, but they did not prove successful in the long run [10].

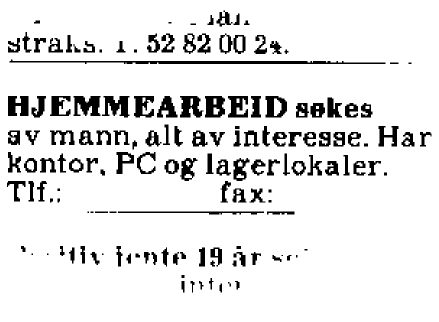

HJEMMEARBEID søkes av mann, alt av interesse. Har kontor, PC og lagerlokaler. Tlf.: fax:

Figure 4 A prospective teleworker seeks employment (Source: Aftenposten)

In spite of few formal teleworking schemes or company policies, home-based teleworking grew to some volume. In the mid-nineties some nine percent of the working population perform parts of their work-related activities at home, and can therefore be called teleworkers in a broad sense of the word [17, 18]. The majority of these teleworkers combine work at home and work at another location. There is a fairly equal distribution among men and women – 60 percent men and 40 percent women – and approximately two thirds of the teleworkers are employed, one third self-employed. The teleworkers represent a multitude of occupations, the more important being education and research, managerial work, office work, sales and accountancy (cf. Figure 3).

The majority of the teleworkers had themselves taken the initiative to establish their teleworking scheme (cf. Figure 4, where a prospective teleworker advertises his services).

In order to bridge the distance between employer and employee, telephony is the most common method, followed, in order, by showing up in person, post, and telefax. Regarding technologies and services deployed in their work as teleworkers, telephony proves to be the undisputed number one, followed by personal computers, the postal system, cellular telephony, answering machines and photo copiers [19].

The study demonstrates differences in communication behaviour between teleworking employees and self-employed teleworkers. In general, the self-employed report a higher level of usage of most types of information- and communications technologies, compared to the employees. A notable exception is telephone conferencing, where the employees have a higher score.

In studies of communication patterns, the *lack* of usage of information- and communications technologies is also of great interest. The survey of teleworking in Norway confirm that a number of teleworkers do not deploy any communication technologies at all. One implication might be to reconsider the conceptions of teleworking. Whereas teleworking is often depicted as a rational usage of recent developments in information- and communication technologies, these findings indicate that teleworking might also be understood as a way of obtaining peace, quiet and concentration, away from the workplace.

In the mid-nineties a teleworking development programme was launched in Norway. This programme was initiated and financed by the Research Council of Norway, with additional support from Telenor Research and Development. The core activity of the programme is a series of pilot experiments in a number of companies and institutions. Further, a number of working groups are established, to discuss and formulate policies on selected themes. The objective of the programme is to contribute to the further development of teleworking, through systemizing teleworking experiences, authoring a teleworking guide, and establishing a teleworking forum [20].

3 Science, good practice, and regulations

> *The role the physical environment plays in influencing organized activity [...] has received only modest attention, particularly in the corporate world.* [21]

A recurrent theme in studies of organizations, is the importance of the physical working environment. This theme is treated in subject areas as diverse as organization theory, architecture, design, and ergonomics. In this section, selected perspectives on the physical working environment will be outlined, as a background for a more detailed discussion of the ergonomic challenges of teleworking. Since ergonomics as a discipline does not cover all substantial areas equally well, in this presentation, fragments from organization theory, architecture and design will provide additional insights.

3.1 The materiality of the workplace

In organization theory, it is customary to regard F.W. Taylor [22] as one of the founding fathers. He performed several studies of work processes, where he systematically manipulated parameters like performance of work tasks and choice of tools, and found substantial productivity impacts through what he called 'scientific management'. The productivity impact was partially due to efficiency improvements, partially due to a reduction of shirking. He recommended changes in the division of labour, where managers should plan the production process, whereupon the workers in detail should follow the plans and instructions [23]. An even more thorough reorganization of the work processes was instituted through the development of the assembly line, usually associated with the name of Henry Ford. Here, work tasks were subdivided into small, repetitive elements, and pacing and control were to a large degree 'delegated' to the speed of the line (this production technology is parodied in Chaplin's film *Modern Times*).

These examples, and numerous more, illustrate how central reflections on the physical working environment – often conceptualized as production technologies – were in the early phases of organization theory (see also [24]). Later, the scope of interest has broadened to include office work and service work, and more detailed themes like strategy formulation, the importance of the informal organization 'within' the formal one, and the role of management.

One of the classic authors of organization theory, Chester I. Barnard, formulated what may be read both as a manifesto and as a research program for the role of the physical environment in organizations. In *The Functions of the Executive*, he distinguishes between on the one hand the terms enterprise, business and operation, and on the other hand the term *organization,* which is "reserved for that part of the cooperative system from which physical environment has been abstracted". This process of abstraction then forms a basis for more detailed investigations of relationships between "all aspects of the physical environment" and the organization, "as may be required for the purpose in hand" [1]. (See also the quotation above.)

Several researchers follow Barnard's methodological dictum, whereas other research traditions present more abstract and disembodied approaches, where organizations may be portrayed as 'nexuses of treaties' [25], or through metaphors like 'brains', 'cultures', or 'political systems' (see the discussion in [26]). Theories of the 'information society', the 'post-industrial society', and the 'post-modern society' [27] may further encourage an understanding of organizations – or entire societies – as 'immaterial' [28]. This development towards a disembodied and immaterialized theory is countered by other approaches, emphasizing the materiality and corporeality of work, organizations and social encounters. The bodily responses of work-related stress and repetitive strain injuries demonstrate clearly that the lightness of information work may be quite unbearable [29].

Workplace architecture and design represent perhaps the most obvious forms of physical working environment, and treaties on factory design and office design provide valuable insights regarding the way task performance, co-operation, and management are sought embodied or materialized in buildings: "Influencing behavior is almost all of what management is about, and buildings influence behavior" [21].

In the development of Ford's assembly line factories, the building designs of architect Albert Kahn were integral elements. The buildings were tailored to the production processes, in particular to the flows of materials and cars under production [30]. These assembly line factories were, however, best suited for long production series of uniform products, since changing the production flow was a time-consuming process.

Office buildings do not present the same building requirements as factories do, since the handling, transportation and storage of information is comparably simple and flexible – even in paper-based offices. Office work may therefore be performed in a variety of settings. Certain types of office work have a long tradition in western societies, like banking, trade and commerce, and the work of lawyers and notaries: "The notary was the inventor of the private office as we know it" [31]. Until the eighteenth century, offices were usually integrated in the private home of the notary, lawyer or trader, later, offices became located in designated office buildings, thereby reinforcing the physical demarcation between the public and the private spheres.

Early office buildings were similar to apartment buildings, but gradually office space became constructed on the factory model, with work in open spaces, and with a supervisor centrally located, overlooking the employees. Later, the office building took the form of cellular offices arranged on each side of long corridors, with position and status manifested in the size of the offices, as well as the dimensions of the desk, type of chair and so on. Subsequently, the office buildings were supplemented with the landscaped open-plan office (or *Bürolandschaft*) where the employees typically were located according to workflow processes, the compartmentalized office, and the flexible office [31, 32].

Around the turn of the century, office work changed radically: Office work became a particular activity in a broadening range of businesses; a feminization of the white collar work force started; and the status of office work deteriorated.[1]

[1] *It goes beyond the scope of this essay to even try to outline the mechanisms interrelating these development trends.*

The principles of scientific management were also applied to office work, with differentiations within the category of office work as one result [31, 32, 33].

From the late nineteenth century on, a number of technological innovations were introduced in the office, like the telephone, the typewriter, carbon paper, and vertical filing systems [34], and specialized office furniture was developed to accommodate an increasingly diverse range of work tasks: Office desks with 'scientifically' designed filing space were developed, as well as the sunken well desk with surfaces at different heights for the typewriter and for writing. With office work as a full-time occupation, not as one work task among others, the design of chairs became an important topic, and adjustable chairs, designed for a multitude of sitting postures – but with different designs for different work statuses – became standard inventory in offices [32].

Whereas organization theory and workplace architecture and design present comprehensive, overall perspectives on the relationships between the physical work environments, work tasks, and organizations, their attention to detail is somewhat scarce and unsystematic. In contrast, *ergonomics* (or human factors, as it is also called) is a workplace science presenting detailed knowledge of work situations and topics like lighting, noise and air quality, as well as working postures, task design, stress and fatigue [35]. Ergonomics as a science does also present a normative orientation towards an improvement of the interaction with the working environment, and is therefore practically oriented: "Ergonomics has little value unless it is applied" [35].

There are many specialities in the field of ergonomics, like physical and cognitive aspects of user-machine interfaces, workplace design and workspace layout, physical and psychological working environment, and job design, selection and training (after [36]). In areas of ergonomics, one may observe a shift in focus from studies of thoroughly designed workplaces like a control room or a cockpit, towards studies of less structured work situations, like office work in general, not solely typing or filing, or the functioning of work groups aided by computers (see [29] and the contributions in [37]).

Although ergonomics appears as a conglomerate of particular studies, without an integrated theoretical perspective, each of the elements of the conglomerate provides valuable insights for workplace design – and thereby also for the design and evaluation of teleworking as a working environment.

3.2 Regulations of the working environment

Ergonomics as a science, as well as organization theory, architecture and design, are all international in their orientation. On the other hand, sets of rules and regulations for the physical working environment are typically anchored in local traditions and national legislation, like the respective worker protection acts and corresponding regulations. In some countries, the area of jurisdiction for worker protection regulations is rather restricted, whereas other countries have widely encompassing regulations. Other national differences may be found in the degree of specification of the respective regulations, and the means instituted to ensure the regulations will be followed (for a study of European approaches to labour legislation, see [38]).

There are certain attempts of creating international regulations of teleworking, like the ILO Convention and Recommendation concerning home work, adopted by the International Labour Conference at its eighty-third session in 1996 [39, 40]. It may, however, still take some time before they will show concrete results, due to the extensive process of ratification and implementation in the respective countries. Within the framework of the European Union, telework has been discussed as one instrument for the further development of a unified European labour market [41].

In Norway, the Worker Protection Act was in 1977 replaced by a Working Environment Act characterized by an encompassing area of application. The objective of the Act is to "secure a working environment which affords the employees full safety against harmful physical and mental influences"[2], according to standards that correspond to "the level of technological and social development of the society at large at any time". Further, it is the objective of the Act to secure "sound contract conditions" and "meaningful occupation for the individual employee", demonstrating a strong normative orientation where the workplace is acknowledged as an arena with mutual social obligations, not simply a meeting place, where work efforts are exchanged for money. Finally, it is the objective of the Act to provide "a basis whereby the enterprises themselves can solve their working environment problems in co-operation with the organizations of employers and employees" (see also [42]). Thereby, a processual approach is chosen, where the legislation constitutes a framework for local activities for the improvement of the working environment. Regulations are formulated under the Act, and the Directorate of Labour Inspection has a responsibility for its enforcement. In spite of its enterprising objectives, the Norwegian Working Environment Act does not apply to "work performed in the employee's home", although regulations may be formulated [43].

4 Ergonomical challenges of teleworking

"The King shall decide whether and to what extent this Act shall be applicable to work performed in the employee's home"

The Working Environment Act

For teleworking to be possible, certain work tasks must be capable of relocation to other whereabouts than the traditional workplace.[3] The crucial importance of information and communication technologies for teleworking, is that this family of technologies allows a vast number of work tasks to be performed at an arbitrary location, and thereby being potential candidates for teleworking. If teleworking solely related to the single work tasks in question, one could argue that it presented few generic working environment challenges, if any: The respective work tasks would present the same challenges as the not-relocated counterpart, and should meet the same requirements, irrespective of the location where the work tasks were performed. This line of reasoning would imply that the ergonomics of teleworking was no different from the ergonomics of the similar work task in another setting or location. While bringing attention to the general issues of

[2] *The passages from the Working Environment Act are quoted from [42].*

[3] *An occupation or a job encompasses numerous work tasks.*

ergonomics to the field of teleworking studies, there is a major flaw in this argument: Teleworking is not solely about a potential re-location of work tasks, but also about a realization of this potential, where work is performed in settings with distinctive qualities, providing distinct working environment challenges.

Teleworking represents a dispersion of the work organization, which, in turn, has implications for the functioning of the organization, like management techniques, relations to customers and colleagues, and relations with family and friends. One may argue that the geographical dispersion of the organization has to rely on a degree of trust between the parties, and will also impose a need for making work results more visible, as in the management technique of 'management by objectives' [7, 14]. An introduction of surveillance methods [44] may, however, turn out to be counterproductive, since surveillance may erode the crucial element of trust. National legislation and collective agreements may also regulate the usage of surveillance methods [43].

The dispersion of the work organization has also implications for the way positions are acquired and manifested in a decentralized and somewhat elusive organization [7, 8]. Teleworking may also have implications for the distribution of work tasks in an organization, since this way of organizing work often implies that the teleworker performs certain work tasks previously taken care of by colleagues, supervisors, and secretarial services.

Teleworking has been the topic for a large number of predictions, recommendations and warnings, but has so far turned out to be less drastically different from 'ordinary work life' than imagined: Teleworking has proved to be more home-based than expected, whereas teleworking centres have not been able to fulfil the promises made; mobile or nomadic teleworking has grown in importance, and has got recognition as one of the main forms of teleworking; and finally, teleworking has not turned out to be a full-time employment for the majority of teleworkers, but an option for a share of the working week, or an additional or supplementary activity.

Teleworking is not a 'pre-packaged' way of organizing work – there are several options and choices to be made, while there are relatively few traditions and regulations for this area of work. Central questions are how teleworking organizations are designed, who are in charge of the design, and the responsibilities involved if the result proves to be an unsatisfactory working environment. Topics like workplace inspection rights, workplace insurance, regulation of working hours, monitoring of efforts and efficiency, and communication of expectations and results are but a few arenas where responsibilities have to be reassigned, and new procedures have to be established.

4.1 Working at home

The basic requirement for teleworking to be possible, is that selected work tasks may be performed at locations other than the employer's. The home may, accordingly, serve as a workplace, given that it fulfils fundamental demands of a working environment – including adequate work space, ventilation, lighting, and correct working postures, as well as safe storage, and access to the necessary information, tools and work materials [45, 46]. A common advice in teleworking guides is to allocate one room to an office to ensure professionality, at least when teleworking is not an accidental or sporadic activity. Another advice is to install a telephone line dedicated to work, both to ensure a professional attitude towards the surroundings, and to facilitate the income tax return. If the teleworker expects customer visits to the office, higher office standards may be required, and the teleworker may even find that zoning regulations will not allow customer visits in residential areas.

Home-based teleworking acquire distinct characteristics through its location in the home: Home and work are two spheres with their own dynamics, but the two spheres are also closely interrelated in a number of ways. Our contemporary society is characterized by a fairly distinct separation, in space and time, between home and work, and designated work spaces and working hours function as major mechanisms for the co-ordination and synchronization of societies. This is contrary to earlier societies, where work was performed at home, without sharply defined working hours. Teleworking, and in particular home-based teleworking, challenges the current distinction between home and work, as the home is deployed also as a work space. This means that when working at home, the teleworker somehow has to manage a boundary between home and work, by transforming or re-establishing what has been a territorial and time-based differentiation, into a functional differentiation between two classes of activities. One strategy is to have one room as a designated work space, maintaining a geographical distinction between work and non-work, while the practice of a strict time-schedule is another strategy for upholding a home/work-boundary [16].

The strategy of practicing standard working hours means that belonging to the sphere of home or work may be determined by the time of the day. This strategy may be instrumental for managing role expectations versus colleagues, customers, family and friends – for instance regarding the timeframes during the day it is appropriate to make or required to receive a phone call. With very flexible working hours, or with work interlaced with periods of non-work, the boundary management becomes even more complex (in particular for the surroundings).

Work proves to be a gendered practice, as is seen in the distribution of men and women in various occupations, their respective positions in bureaucracies and corporate hierarchies, and in the distribution of wealth and obligations in their work careers [47, 48]. Detailed studies of businesses prove that gender differences at work change over time, and that numerous actors and activities both uphold and change gendered perceptions and practices [49]. The home is another thoroughly gendered arena, and one may presume that teleworking, and in particular home-based teleworking, will be understood differently by women and men. One conjecture may be that for male teleworkers, work will permeate the home sphere, while for female teleworkers, it will be the other way round.[4] A Norwegian study of motivations for becoming teleworkers support this hypothesis, as a main motivation for male teleworkers was that they did not manage to complete their work tasks during ordinary working hours, while for the female teleworkers, a main motivation was a wish to take better care of family and children [18].

The teleworker does not only have to uphold and manage a distinction between

[4] *I owe this formulation to Øystein Gullvåg Holter.*

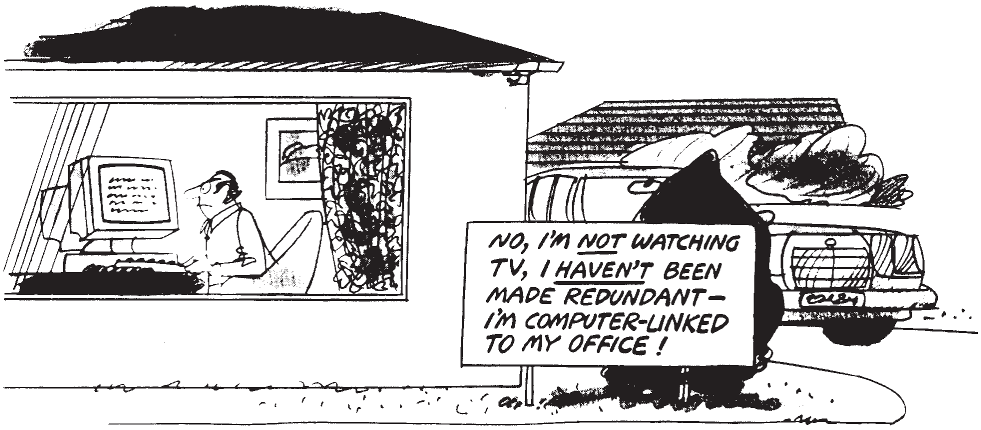


Figure 5 Teleworking arrangements may encounter legitimation crises (Source: [50])

the spheres of work and non-work, but must also manage the transition from one sphere to the other. For ordinary work, the journey between home and work may serve as a very concrete *rite de passage*, whereas for telework, more symbolic rites may be appropriate. In the literature, one finds advises like following the ordinary workplace rhythms, or dressing as if going to a workplace outside the home.

4.2 Working in a not fully equipped office

In the last few years, computers have become a fairly common household appliance, as the prices of computer equipment have fallen sharply, and numerous applications and services have motivated the purchase of a home computer. In some respects, the teleworking office may therefore be as well equipped as an ordinary workplace for the core activities or tasks of the teleworker. It seems, however, to be a fair assumption that few teleworking offices provide the whole range of welcome equipment of a well established company, in particular in the case of mobile or nomadic teleworking. Required equipment may be almost everything from a first-rate copying machine, a colour laser printer, a shredder, or an able stapler. Further, the teleworking office may suffer from the absence of secretarial services, postal services, the services of a receptionist, or a computer service department – all factors which may influence work task quality and workplace performance.

If the teleworking activities are severely affected by a spartan working environment, one solution is to purchase the respective services from an independent provider, like a printing or copying service, or a third party computer service. Other solutions may be relevant, if the ascetically equipped teleworking office is not the full-time workplace. With additional access to a fully equipped work environment, some of the challenges mentioned may be solved by advancing or postponing certain equipment- or service-dependent activities to the time of the day or the week when the appropriate technology or service is available.

4.3 Working at multiple work sites

A common assumption in several treatments of teleworking, in particular in earlier writings discussing teleworking as a potential or a threat, is that the teleworkers would spend the entire working week in a remote setting. Studies of actual teleworking behaviour prove, however, that the full-time teleworkers represent only a small share of the entire teleworking population, as the majority of teleworkers combines teleworking with work at another site [17].

Working at more than one location may solve certain problems sketched above, regarding not fully equipped offices. On the other hand, this solution requires the teleworker both to maintain an acceptable working environment at each location, and to manage the transitions from work site to work site. The first mentioned concern may be solved by, as far as possible, reproducing the working environment at each (tele)working site. This choice may provide both functional and familiar working environments at each and every site, but may also prove to be both expensive and space consuming. Further, this choice may neglect the wish for a functional differentiation between the various work sites: The mobile teleworker, for instance, may want to have a spartan workplace to allow for mobility. The mobile workplace may then be supplemented by a better equipped, stationary base. Although the strategy of replicating an entire working environment from site to site may not be desirable, access to, or the replication of selected elements, like databases, e-mail accounts or computer programs, may ensure functionality and flexibility in the movement from one site to another. Working at multiple work sites seems to require an increased amount of planning and co-ordination to ensure productivity, accessibility and availability. Working at multiple work sites may also preclude the practice of deploying a personalized workspace, like a shelf or a desktop, as a memory aid and a work instrument.

The practice of working at multiple work sites may also serve as a motivation for not equipping any of the worksites fully, by letting the various workplaces *in sum* provide a fully functional working environment. This is perhaps most evident in the so-called flexible offices, where a functional differentiation between workspaces is expected, and where main offices function more as meeting places or ‘touch-down’-offices (typically with fewer workplaces than employees), in combination with home-based work and work at the premises of customers or clients [13, 51].

4.4 Working in a dispersed organization

Working in an organization does not necessarily imply close co-operation with colleagues. A number of work tasks are performed in relative autonomy and isolation from the colleagues, like contact with customers or concentration-intensive work. In such cases, the geographical dispersion of the organization introduced by teleworking, does not have to present major challenges for the performance of the singular work tasks. In the teleworking literature, there is a common observation that work task independence is positively correlated with teleworkability [7, 52]. This observation may, in turn, be taken as a *recommendation* for a decoupling of work tasks, in order to allow for teleworking. This recommendation may, however, be contrary to the successful realization of the organization's goals, or detrimental to the functioning of the organization as a social system, with consequences for organizational learning, and organizational belonging. One may therefore argue that this approach is rather short-sighted, by considering telework or teleworkability a goal in itself, not as means for improving the quality of work and the quality of life.

Work tasks relying on extended co-operation with colleagues will typically be more profoundly affected by a geographical dispersion, since co-operative practices, as well as conceptions of co-operation, typically rely on an assumption of co-location. Distributed co-operative work presents challenges for meetings and discussions, for the distribution and access to information, for the co-ordination of activities, for organizational learning and the formation of shared opinions, and for joint access to the work material and work tasks in question. In the vivid research areas of computer-supported co-operative work (CSCW) and distributed work, these challenges are discussed, and technologies for solving them are developed like document management systems and screen-sharing and joint editing technologies (see the readings in [53] and [54], among several other sources).

Numerous aspects of co-operation and socialization in an organization are affected by teleworking, as physical location and other physical clues provide valuable information for the functioning of an organization. The importance of this kind of information is acknowledged when considering the multitude of actions being performed during an ordinary work day: Proximity of workplaces is usually an indication of co-operation or belonging to the same work-group, light in an office or a coat on a hanger may indicate that a person is present, an open door indicates availability, while the sound of a conversation, even through a closed door, may indicate that a person is busy and not to be disturbed. In a distributed organization, other clues have to be interpreted, and other procedures have to be established to ensure the smooth functioning of the organization.

4.5 Meanwhile, back at the office

In the teleworking literature, one finds elaborate discussions of promises and potential pitfalls for actual and prospective teleworkers and managers, together with advises and guidelines for the successful implementation and practice of telework. Teleworking studies do, however, in less detail address how the introduction of a teleworking arrangement affects those in an organization who are not part of the teleworker-manager dyads. Those remaining at a common office, either by choice or by not being allowed to telework, will be affected by a teleworking arrangement in several ways: Their work tasks may change as a complement to the changes introduced at the teleworking sites, as discussed above. Further, they may have to compensate for, or cover work tasks which otherwise would have been performed by the teleworker if she or he had been physically present, like spontaneous meetings or certain enquiries from customers or clients. With colleagues widely scattered, the in-house workers may also be affected in a similar way as the teleworkers, regarding the changed framework for knowledge development and -transfer, workplace socialization, and the development of company affiliation and company loyalty. Finally, the social atmosphere at a workplace may be affected by teleworking, in particular when there is an extended telework programme, with only a small share of the employees present at a given time.

4.6 Workplace symbolism

The workplace is a site both for productive work and for the production and circulation of meanings and interpretations, inseparable from 'productive work'. *All* economic activities are inherently cultural phenomena, because "the processes of organization and production [...] are the subject of cultural change and reconstruction [and] ideas, images values provide us with shared frameworks, assumptions and moralities which we use to define and make sense of our work and employment" [55]. The cultural embeddedness of an organization is explicitly discussed in the literature on organizational culture, where culture is acknowledged as the 'glue' upholding the organization as a functioning, productive unity: "The real business goes on in the cultural network" [56].[5]

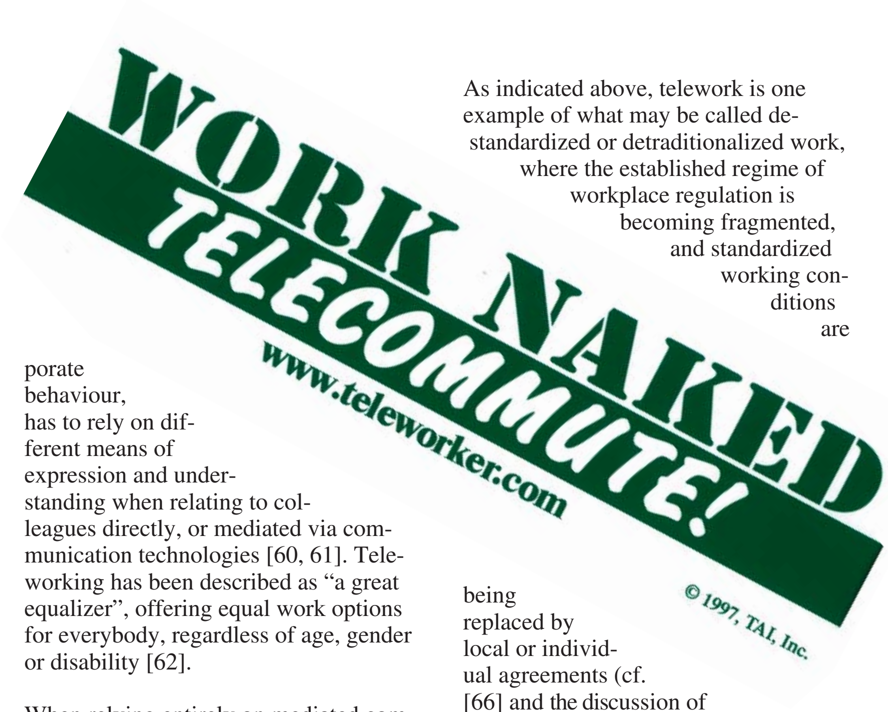
Corporate architecture is a major arena for workplace symbolism (see the contributions in [57]). It is common for major corporations to erect centrally located office buildings as signs of importance and wealth. Designated buildings and canteens for 'white collar' and 'blue collar' workers serve as instruments for in-group communication and identification, and for upholding status hierarchies. Further, the size and location of offices are common indicators of status in the corporate hierarchy, an anteroom with a secretary demonstrates power, and the possibility to refurbish and personalize an office is an acknowledgement of one's importance for the organization (see the discussion of 'power offices' in [31]).

With teleworking, the corresponding activities and processes will follow somewhat different rules than at a co-localized workplace, in particular when a substantial portion of the working week is spent outside a common work site, or when a flexible office solution with 'non-territorial' offices has been adopted. While for instance career development may be manifested in the physical relocation from a smaller to a larger office, or from one wing of a building to another – in a telework setting, other markers of group affiliation and social status have to be developed.

For corporate success, deliberate 'impression management' is recommended, including proper clothing, politeness and attentiveness, and non-verbal impression management [58]. In a geographically dispersed organization, the 'presentation of self in everyday life' will follow different rules, than "within the physical confines of a building or a plant" [59]. Impression management, and the corresponding interpretations of cor-

5 *Here, 'culture' is typically understood as a variable, subject to manipulation in the pursuit of organizational productivity (for a discussion and critique, see [55]).*

porate behaviour, has to rely on different means of expression and understanding when relating to colleagues directly, or mediated via communication technologies [60, 61]. Teleworking has been described as "a great equalizer", offering equal work options for everybody, regardless of age, gender or disability [62].



When relying entirely on mediated communication, it is possible to stage the presentation of one's self rather freely, as is epitomized in the *New Yorker* cartoon "On the Internet, nobody knows you're a dog" (quoted in [63]) or in the car bumper sticker text: *Work naked: Telecommute!*

4.7 Workplace design and workplace democracy

Portability of work tasks is a prerequisite for teleworking to be possible, but teleworking is not solely about work task performance at an arbitrary location, as telework also relates to the frameworks within which work tasks are performed, like the contractual arrangement between the teleworker and the employer, or the teleworker's influence on his or her working environment and working conditions. All work is embedded in frameworks of labour legislation, agreements (collective and/or individual), traditions and norms. For 'ordinary work', this framework is fairly extensive, whereas for telework, the corresponding set of regulations is poorly developed.

Teleworking has by some authors been seen as an opportunity for a radical reorganization of working life: "[T]elecommuting could be that bit of reasoning that will allow managers to once again take control without the use of a whip and bullhorn" *(sic)* [64]. This type of quotations demonstrates clearly the importance of addressing normative issues regarding the future development of teleworking [65].

As indicated above, telework is one example of what may be called destandardized or detraditionalized work, where the established regime of workplace regulation is becoming fragmented, and standardized working conditions are being replaced by local or individual agreements (cf. [66] and the discussion of modernization processes in [67]).

The absence (relatively speaking) of traditions and regulations requires a number of choices to be made by the parties involved. With several degrees of freedom, few paradigmatic models, and few standards and regulations, the prospective teleworkers or teleworkers' employers may face what may be seen as an insurmountable complexity of choices, or they may settle for solutions inappropriate for their organizational needs. The many options and choices in teleworking also actualize questions about how, and by whom, a corresponding framework should be designed, to what extent this framework should be detailed, and for whom it should be applicable [39, 43].

As recommended in the handbooks and guidelines, several teleworking programmes start with a pilot project to try out technological and organizational solutions and to develop a contractual framework for teleworking. Typically, the next steps will be an evaluation of the pilot project, and a decision regarding the future of teleworking in the organization. Advancing from a pilot project towards full-scale teleworking, one encounters questions of the scale and scope of teleworking in an organization: Is teleworking for the selected few, just like an additional benefit, or shall the majority of the employees be eligible as candidates for teleworking.[6] When extending the basis of potential teleworkers, one also encounters a diversity of motivations, competences, and need for support.

Teleworking represents challenges for the way workplace influence, in all its forms, is performed, due to the geographical dispersion of the work organization.[7] Further, the geographical dispersion has implications for how opinions about workplace policies are formed, a factor of major importance for the practice of workplace democracy, usually taken care of in informal meetings around coffee machines, water coolers, and luncheon tables.

5 Implications for future research

> *The advent of modernity increasingly tears space away from place by fostering relations between 'absent' others, locationally distant from any given situation of face-to-face interaction.*
>
> [67]

A main objective of this presentation has been to elaborate certain working environment challenges of teleworking. As a result, the presentation may be read as a catalogue of problematics and stumbling blocks for the design of dispersed work arrangements, or even as an argument against distributed work. A problem-orientation is, however, inherent in discussions of ergonomical challenges, whereas the pleasures and productivity impacts of teleworking belong to another framework of discussion.

A more appropriate reading of this presentation is that teleworking is to be understood in contrast to, or as an amendment to, specific work practices and working life institutions, and that the ergonomical challenges of telework have to be compared with the ergonomics of existing work life situations. Teleworking as a relatively new way of organizing work, has to prove its relevance and attractiveness, while existing work life arrangements have an authority through their sheer existence and familiarity.

6 *Here, one may add the question: Will having a home suitable for teleworking become a requirement for being employable?*

7 *In this respect, teleworking is quite similar to shift-work.*

There are numerous technological and organizational responses to the ergonomical challenges of teleworking outlined in this presentation, like the development and introduction of appropriate information and communication technologies [68], the establishment of communication practices and electronic meeting places equally accessible for everybody in an organization, or the introduction of meeting days where everybody is expected to be physically present [69]. It has not, however, been the intention in this presentation to outline potential solutions in any detail.

With teleworking as a fairly new way of organizing work, there is an abundance of topics for further research. It is of great importance to systematize existing research, to monitor teleworking practices and experiences – with special attention to gendered practices and experiences – and to raise normative issues to contribute to the further development of teleworking as an attractive way of organizing work. In particular, there are research challenges in following teleworking arrangements beyond pilot experiments, since one may assume a ‘honeymoon effect’ in the initial phases. One may also assume that experiences will differ, when teleworking becomes an option for others in an organization than the most enthusiastic teleworking pioneers. Further, there are important research topics regarding the development of even more ‘elusive’ organizations, with distributed work as a dominant work practice.

5.1 Developing teleworking regulations, practices and concepts

As pointed out above, the phenomenon of teleworking is somewhat ill-defined, and it has been proposed to abolish the term completely. The concept has, however, its acknowledged benefits: “Despite its lack of precision, the word telework has acquired a potent symbolic value” [8]. The problems with establishing a clear-cut definition of teleworking may be lamented as an indication of the immature status of the teleworking research, but may equally well be taken as an indication of the multitude of choices in the establishment of teleworking practices – concordant with the characterization of teleworking as a destandardized or detraditionalized work practice. Hence, the other side of the (alleged) immature status of teleworking and teleworking studies is the privileged opportunity for studying a new organizational form in the making.

The establishment of a regulatory framework for teleworking will impose certain constraints on the multitude of choices in teleworking. The development of dominant practices through a combination of evolutionary processes and the impact of teleworking handbooks and teleworking development programmes will in a similar way, although far less coercive, serve to make the number of choices less overwhelming. There are important research tasks in the monitoring of these emergent practices, in the evaluation of teleworking programmes [20], and in a conceptual differentiation between various forms of what today is called teleworking.

5.2 The ergonomics of teleworking and the future of work

Ergonomics as a discipline is born from a concern for the body in worklife situations, where unsatisfactory working conditions are met with the ‘revenge of the body’ through work-related illnesses, and a reduction of productivity and a deterioration of the quality of work life. Corporeality is also central in studies of culture and communication emphasize the body as a locus for intention and meaning, and the richness of face-to-face communication: Wherever work is performed, one does take one’s body to work [60, 61, 70].

Further studies have to be made about the usage of mediated communication in teleworking and other distributed worklife situations, of media choices, and finally, about the development of communication technologies, and the appropriation or ‘domestication’ of these technologies through the development of appropriate communication strategies [71].

A predominant theme in this presentation has been the malleability of work arrangements, which has been illustrated through glimpses from the history of work, and, in more detail, through a presentation of the numerous choices to be made for a ‘detraditionalized’ way of organizing work like teleworking. Choices represent both potentials and burdens, and the lack of regulations and paradigmatic examples of teleworking means that the prospective teleworking organization is forced to spend some effort specifying the work arrangement.

An interest in the ergonomics of teleworking draws attention to the materiality of work life, like the roles of technologies and workspace architecture for co-operation and work task performance, and serves as a theoretical impetus for acknowledging the importance of the minutiae of work life practices.

The malleability of work arrangements is by no means confined to the establishment of teleworking schemes – all work arrangements are results of a multitude of choices and decisions, but for existing work arrangements they are rarely as open-ended and explicitly discussed. In the on-going transformation of economic life, one finds that every aspect of work is increasingly more exposed to deliberation and competition, and work tasks may be re-configured, re-assembled, or obliterated, and where teleworking is but one of many workplace and market innovations (see the contributions in [72, 73]). Teleworking is in this respect interesting, since it challenges conceptions of the localization of work, and proves that work and non-work are not two entirely separable worlds. Further, when comparing with other ways of organizing work, teleworking demonstrates most clearly the importance of communication in organizations, with the associated competences, cultures and technologies.

6 References

1 Barnard, C I. *The functions of the executive.* Cambridge, Harvard University Press, 1968.

2 Nilles, J et al. *The telecommunications-transportation trade-off.* New York, John Wiley, 1976.

3 Kraut, R E (ed.). *Technology and the transformation of white-collar work.* Englewood Cliffs, Larence Erlbaum, 1987.

4 Toffler, A. *The third wave.* New York, Pan Books, 1981.

5 Julsrud, T E. Teleworking – the vision. *Telektronikk,* 92(1), 16–25, 1996.

6 Negroponte, N. *Being digital.* London, Hodder & Stoughton, 1995.

7 Nilles, J. *Making telecommuting happen.* New York, Van Nostrand Reinhold, 1994.

8 Huws, U et al. *Telework : Towards the elusive office.* Chichester, John Wiley, 1991.

9 Bakke, J W. *Å telle fjernarbeidere.* Kjeller, Telenor Research & Development, 1996. Unpublished document.

10 Hetland, P et al. *Nært, men likevel fjernt.* Stavanger, Rogalandsforskning, 1989.

11 Spinks, W. Satellite and resort offices in Japan. *Transportation,* 18, (4), 1991, 343–363.

12 Club de Bruxelles. *The future of the information society (the 'Bangemann Report').* Brussels, 1994.

13 Becker, F, Steele, F. *Workplace by design.* San Francisco, Jossey-Bass, 1995.

14 Kugelmass, J. *Telecommuting : A manager's guide to flexible work arrangements.* NY, Lexington Books, 1995.

15 NOU. *Telematikk : Teleutvalgets utredning II.* Oslo, Universitetsforlaget, 1983. (1983:32.)

16 Nippert-Eng, C E. *Home and work.* Chicago, The University of Chicago Press, 1996.

17 Bakke, J W. Fjernarbeid – en visjon i forandring. In: J Moe (ed.). *Fra Kardemomme by til cyberspace. Arbeidslivets utfordringer og muligheter i informasjonssamfunnet.* Trondheim, Norges Tekniske Vitenskapsakademi, 1995.

18 Bakke, J W, Bergersen, E. *Teleworking : The Norwegian experience.* Kjeller, Telenor Research & Development, 1996.

19 Bakke, J W. Communication patterns in teleworking. In: K Nordby et al. (eds.). *Human-computer interaction.* London, Chapman & Hall, 1995.

20 Bakke, J W et al. *Teleworking development programme in Norway.* Kjeller, Telenor Research & Development, 1997.

21 Seiler, J A. Architecture at work. *Harvard Business Review,* Sep-Oct, 111–120, 1984.

22 Taylor, F W. *The principles of scientific management.* NY, Norton, 1967.

23 Rose, M. *Industrial behaviour.* Harmondsworth, Penguin, 1982.

24 Rabinbach, A. *The human motor.* Berkeley, University of California Press, 1990.

25 Aoki, M et al. (eds.). *The firm as a nexus of treaties.* London, Sage, 1990.

26 Morgan, G. *Images of organization.* London, Sage, 1986.

27 Kumar, K. *From post-industrial to post-modern society.* Oxford, Blackwell, 1995.

28 Diani, M (ed.). *The immaterial society.* Englewood Cliffs, Prentice Hall, 1992.

29 Grandjean, E. *Ergonomics in computerized offices.* London, Taylor & Francis, 1987.

30 Brunnström, L. *Den rationella fabriken.* Umeå, Dokuma, 1990.

31 Pélegrin-Genel, E. *The office.* Paris, Flammarion, 1996.

32 Forty, A. *Objects of desire.* London, Thames and Hudson, 1986.

33 Lupton, E. *Mechanical brides.* NY, Cooper-Hewitt National Museum of Design, 1993.

34 Yates, J. *Control through communication.* Baltimore, The Johns Hopkins University Press, 1993.

35 Wilson, J R, Corlett, E N (eds.). *Evaluation of human work.* London, Taylor & Francis, 1995.

36 Galer, I (ed.). *Applied ergonomics handbook.* London, Butterworths, 1987.

37 Baecker, R M et al. (eds.). *Readings in human-computer interaction.* San Francisco, Morgan Kaufmann Publ., 1995.

38 Nyström, B. *EG och arbetsrätten.* Stockholm, Publica, 1992.

39 ILO. *Home work. Report V (1). International Labour Conference 82nd Session 1995.* Geneva, International Labour Office, 1994.

40 Haakonsen, H. Taus arbeidsgiverprotest. *Arbeidsliv,* No. 11, 4–5, 1996.

41 *Actions for the stimulation of transborder telework and research cooperation in Europe.* Brussels, European Commission DG XIII, 1996.

42 Gustavsen, B, Hunnius, G. *New patterns of work reform.* Oslo, Universitetsforlaget, 1981.

43 Jakhelln, H. Fjernarbeid. In: H Jakhelln and J W Bakke (eds.). *Fjernarbeid og arbeidsrett.* Kjeller, Telenor Research & Development, 1996.

44 Booth, W. Big brother is counting your keystrokes. In: S R Corman et al. (eds.). *Foundations of organizational communication : A reader.* NY, Longman, 75–76, 1990.

45 Rosenbaum, A. *The complete home office.* Harmondsworth, Viking Studio Books, 1995.

46 Kanarek, L. *Organizing your home office for success.* Harmondsworth, Plume, 1993.

47 Mills, A J, Tancred, P. (eds.). *Gendering organizational analysis.* Newbury Park, Sage, 1992.

48 Kvande, E, Rasmussen, B. *Nye kvinneliv : Kvinner i menns organisasjoner.* Oslo, AdNotam, 1990.

49 Kwolek-Folland, A. *Engendering business : Men and women in the corporate office 1870-1930.* Baltimore, Johns Hopkins University Press, 1994.

50 Pursell, C. *White heat. People and technology.* London, BBC Books, 1994.

51 Raymond, S, Cunliffe, R. *Tomorrow's offices.* London, E & FN Spon, 1997.

52 Shaw, L. *Telecommute! Go to work without leaving home.* NY, John Wiley, 1996.

53 Greif, I (ed.). *Computer supported cooperative work : A book of readings.* San Mateo, Morgan Kaufman, 1988.

54 Engström, Y, Middleton, D (eds.). *Cognition and communication at work.* Cambridge, Cambridge University Press, 1996.

55 Salaman, G. Culturing production. In: P Du Gay (ed.). *Production of culture / cultures of production.* London, Sage, 1997.

56 Deal, T, Kennedy, A. *Corporate cultures : The rites and rituals of corporate life.* Harmondsworth, Penguin, 1988.

57 Gagliardi, P (ed.). *Symbols and artifacts : Views of the corporate landscape.* Berlin, Walter de Gruyter, 1990.

58 Rosenfeld, P et al. *Impression management in organizations.* London, Routledge, 1995.

59 Goffman, E. *The presentation of self in everyday life.* Harmondsworth, Penguin, 1990.

60 Barry, D, Hazen, M A. Do you take your body to work? In: D M Boje et al. (eds.). *Postmodern management and organization theory.* Thousand Oaks, Sage, 1996.

61 Stohl, C. *Organizational communication.* Thousand Oaks, Sage, 1995.

62 Ingham, C. *New work options.* London, Thorsons, 1996.

63 Turkle, S. *Life on the screen.* London, Phoenix, 1997.

64 Mahfood, P E. *Homework . How to hire, manage and monitor employees who work at home.* Chicago, Probus, 1992.

65 Bakke, J W. Telework, participation and workplace democracy. In: P Jackson and J van der Wielen (eds.). *Amsterdam '97. Second international workshop on telework.* Tilburg, Work and Organization Research Centre, 1997.

66 Beck, U. *Risk society.* London, Sage, 1992.

67 Giddens, A. *The consequences of modernity.* Cambridge, Polity Press, 1990.

68 Eldib, O E, Minoli, D. *Telecommuting.* Boston, Artech House, 1995.

69 Sakshaug, U et al. *Håndbok i fjernarbeid.* Oslo, Norges forskningsråd, 1997.

70 Turner, B. *The body & society.* London, Sage, 1996.

71 Haddon, L, Silverstone, R. *Teleworking in the 1990s : A view from the home.* Brighton, University of Sussex, 1993.

72 Olberg, D (ed.). *Endringer i arbeidslivets organisering.* Oslo, FAFO, 1995.

73 Hesselbein, F et al. (eds.). *The organization of the future.* San Francisco, Jossey-Bass, 1997.

John Willy Bakke is Research Scientist at Telenor Research & Development, Kjeller.
He is currently completing his doctorate at the Department of Sociology and Human Geography, University of Oslo.

e-mail:
john.bakke@fou.telenor.no

Human factors of videotelephony

KNUT NORDBY

Experience with videotelephony in Norway and elsewhere has clearly shown the need for operating procedures and controls to be as simple and easy to use as possible. It was decided to incorporate human factor aspects to ensure a high degree of usability from the very beginning.

A videotelephone is something much more than just a sound telephone with a video picture added. In the videotelephone we have to adopt new ways of communicating with other people and we must learn to operate several new controls, the results of which are not always obvious to us. This poses many new and different problems from the point of view of human factors compared to the much simpler sound telephony. This article is based both on well established and accepted practices in the fields of cognitive and human factors and on experience gained in our trials.

1 Visual distance communication

To be able to see and talk to a person who is not present is one of man's oldest dreams, the beginnings of which go from way back to the mist of mythology and fairy tales. Reflections in fluids, crystal balls and magical mirrors (as the mirror on the wall in Snow White) are examples of fanciful metaphors for 'audio-visual' distance communication which have become part of our folklore.

Figure 1 Alleviating absence by use of the 'Telephonoscope'. A mirror is envisaged as the display screen. (From Albert Robida 'Le Vingtième Siècle'*, 1880s)*

Even as late as the 1880s, the French futurist writer Albert Robida in his book *"Le Vingtième Siècle"* envisaged a mirror on the wall as the display for his 'Telephonoscope'. This metaphor was also used by Jules Verne for the 'cinema telephone', a combined sound telephone and 'tele-cinema', in his 1889 short story *"In the Twenty-ninth Century: The Day of an American Journalist in 2889"*.

The possibility to speak to someone not present had been realised in 1876 when Bell and Gray invented the telephone, which soon became an accepted form of communication. It was not until 1927, however, that Herbert I. Ives at the Bell Laboratories made the first telephone call with a picture between Washington and New York, although a picture was only transmitted one way. Technology had proceeded so far that it was now possible to see the person with whom you were talking as well.

Finally, in 1936 in Germany, full two-way picture and audio communication was established between the main post offices in Berlin and Leipzig [1], demonstrating the feasibility of picture telephony, or to use the accepted modern term, videotelephony.

Lately, there have been a number of videotelephones and videotelephone projects around the world, with rather varied results, but no videotelephone has had any commercial success so far. It is very easy to attribute this to such factors as cost, availability, usability and appropriate applications. The topic here is to take a closer look at some important usability factors.

1.1 Basic human factors

Very complex machines – and the telephone system is the world's largest and most complex machine – must have simple basic operating procedures which are consistent with human physical, perceptual and mental capabilities [2]. This is essential if the end-user is to be able to interact confidently and directly with the equipment.

In the beginning, all telephone calls were handled by operators and the only thing the caller had to do was to lift the receiver off-hook and ask to be connected to the desired party. As telephone technology developed, it was necessary to adopt new call set-up procedures, such as dialling and pushing buttons (at first only for local calls, later also for long distance and international calls), but even these were not beyond the capabilities of most seven year olds. Today, however, there is a host of new services which may be accessed from a telephone, but they involve struggling with long numbers, arbitrary service codes and cryptic line signals.

The advanced technology of today's telephony, and other areas, has resulted in increased user-problems and has raised many interesting human factors issues. Although these issues have been dealt with in more detail elsewhere, they hold important implications for videotelephony.

1.2 Knowledge in the head and in the world

To use any telecommunication terminal or system, the user must have a conceptual model of it and of the service it provides. The model must be simple to comprehend and it must cover the essential features of the service. Some models are derived from explicit knowledge in the world, provided by the device itself, by the instructions and by the physical constraints. Other mental models are derived from knowledge in the head and are created in our minds from our interactive strategies and general rules for interaction, analogies, general knowledge, past experience with similar devices, etc.

All control procedures must be consistent with this conceptual model and all the visible and audible parts of the terminal must reflect the terminal's current state in a way that is consistent with the conceptual model.

The user needs simple, intuitively understood and easily remembered mental models, maps and metaphors of what the terminal can do. Thus, the conceptual model does not have to be an accurate representation of what *actually* goes on inside the terminal. It is the end result of the service provided by the system that is of interest, not how it is performed by the system: e.g. it is irrelevant to the user how the picture is transmitted, as long as the picture is displayed on the screen. The mental model or the metaphor here is simply that of a 'distance mirror'.

Very often there is a clear discrepancy between the conceptual model of the designer (if there ever was one) and the conceptual model formed by the user. The designer is in a totally different position from the user since the designer usually has a number of constraints governing the design work: economy, deadlines and tight schedules, compatibility with existing systems, design traditions, corporate identity, marketing department requests, current fashions and trends, cultural traditions, a too detailed knowledge of how the thing actually works, etc.

The user, on the other hand, will have mental models that are based on: realistic or unrealistic expectations of what the device should do, previous experience with similar devices, no detailed knowledge of what goes on inside, sales talk and brochures, myths and rumours, or just wishful thinking. It is important to determine the user's conceptual model of the product at the earliest possible stage in the design process and design to comply with it. The domestic video recorder (VCR) is a good example of a particularly poorly designed product user interface which has given problems to most users.

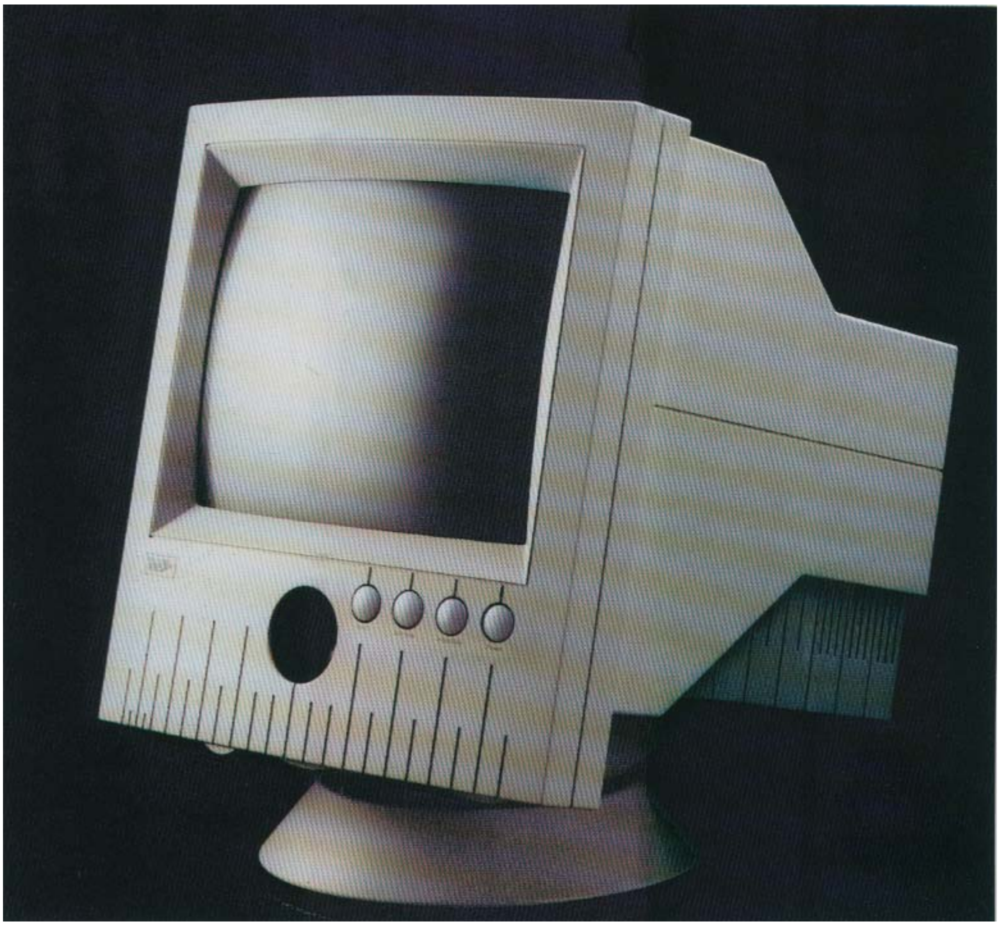


Figure 2 The first design concept for a fully working videotelephone from ALCATEL developed in co-operation with Norwegian Telecom Research (now Telenor R&D) in 1987. Observe the camera lens under the display screen. This position was chosen to improve eye-contact

1.3 Simplify the task

Complex tasks requiring many control procedures are difficult to manage and remember by most people. Wherever possible, tasks should be simplified by changing the structure and minimising the need for problem solving and planning ahead. No demand should be made on short-term memory as this is a limiting factor with most people. Technology should be used to make things visible, relieve short-term memory and give feedback.

Simple in this context means that procedures should be; intuitive (i.e. the natural thing to do); consistent (the same action always gives the same result); easy to remember by infrequent users; easy to learn by first time users; self-explanatory; require little or no instruction; and should give clearly visible results. These requirements are difficult to fulfil on many complex terminals, but they should constitute a design goal to strive for.

1.4 Make things visible

The user cannot sense directly what goes on inside electronic equipment, unless the equipment displays information about its internal state in a form that can be sensed by human observers. To be able to use modern electronic equipment (e.g. videotelephones, computers, etc.) it should not be necessary to understand their internal workings.

Unfortunately, many telephone services, both those offered in the public network and in private business exchanges, make use of cryptic procedures that are difficult to remember. To initiate a service you first have to dial the exchange, then a '*' to initiate the command sequence, then an arbitrary service code number (e.g. '63'), then another telephone number or a time and date, and finish off the sequence by dialling '#'. Normally, no feedback is given during the process and as long as valid numbers are entered there is nothing to warn you of a mistake. If, for example, you want to call the wake-up service to catch a flight, but by mistake you dial the code for diverting incoming calls to another telephone (there is no feedback to warn you) or, having dialled the correct service code you misdial the time or the date, there is no feedback here either. In other words; you have no way of knowing if you will be awakened at the right time – or even the right day.

There is a sinister anecdote told about a man who wanted to make an external call from a strange office he visited. Assuming that an '0' would get him an external connection (in fact it was '9'), he dialled the number. By pure coincidence, the number he dialled reprogrammed the system to broadcast to all the telephones in

Figure 3 The latest Tandberg ISDN portable videotelephone with folding flat display screen, the Tandberg Vision 600

the system and also to turn them all to loudspeaking mode. There was nothing on his terminal to indicate to him that every sound he made was heard in every office in the building.

1.5 Cognitive maps, models and metaphors

In order to understand how to use equipment and procedures, we often devise our own internal "maps" and "models", often naïve and simplified, according to our perception of how the equipment and procedures work. Such mental maps and models are necessary when systems become very large and complex.

Most mental maps and models are based on simple *metaphors* that should be easy to understand and remember. A metaphor is an intuitively understood analogy or parallel from everyday life of a function or process. Some metaphors just naturally come to mind, while others are specifically created to help the user understand and use a complex process or product.

When making new metaphors, it is important that they mimic an essential feature of the process in such a way that it supports the user's naïve concept of the process.

It is essential that metaphors and mental maps make the user understand the relationships between intentions and possible actions; between any actions and their effects on the system; between the actual system state and what the user perceives by the visual, auditory and haptic senses; and between the perceived system states and the user's intentions, needs and expectations concerning the equipment.

Some metaphors are successful in conveying the essential idea to the user, which reduces mistakes and shortens the time to learn a procedure. But metaphors may also be detrimental to the user, increasing the number of mistakes made and the time it takes to learn the process. Metaphors can be psychologically unsound and counterproductive in some applications (i.e. the bin or waste-paper basket metaphor of Macintosh to eject diskettes – the last thing you want to do is to throw your newly made back-up copy in the bin).

1.6 Constraints

One very important way to achieve a simple and user-friendly design is to utilise natural and artificial constraints. The user should get the feeling that there is only one possible thing to do – i.e. the right thing.

A natural constraint is based on what the user *believes* to be the only natural and logical course of action, i.e. when there is only *one* light switch for the room lights, when there is only *one* way out, when there is only *one* button to press, etc.

Artificial constraints means that the device is designed so that one action cannot be executed before another, essential, action has been carried out, e.g. you cannot dial before you have a dialling tone and you cannot get the dialling tone before you lift the handset off-hook, etc.

1.7 Errors

If errors can be made, people will invariably make them. All terminal equipment and systems should be designed to support rather than fight the user. Never blame the user. Nothing is achieved by harassing the user who may already be painfully aware that an error has been made. If errors are easy to make, the designer is nearly always at fault. Design to avoid errors, but expect that errors will always be made and design to make it easy to correct them.

Errors are costly in all systems. They waste time, money and resources, they can endanger people's health and life, and they are costly to correct. To avoid errors through inherent sound design should be the prime concern in all design.

1.8 When all else fails, standardise

In an ideal world, the specific design principles discussed above should help us to overcome most user-problems and make all users perform perfectly all the time. Unfortunately, since we do not live in an ideal world, this is seldom the case.

When arbitrary mapping, arbitrary codes and unnatural operating procedures cannot be avoided, the only way may be to standardise the operating procedures, instructions, controls, signals, messages, prompts, etc. to ascertain that users will be able to learn and use the equipment and the services it provides efficiently. By having universal standards we may be able to learn and use inconsistent and incomprehensible procedures.

1.9 To sum up

All the ideas discussed so far can best be summed up in the two statements made by Donald Norman [2]:

1. Make sure that the user can figure out what to do.
2. Make sure that the user can tell what is going on.

If these two basic recommendations are always followed, it is difficult not to design a workable user-interface.

2 Basic communication abilities

Like ordinary telephony, videotelephony is supposed to make use of our natural communication abilities – i.e. the face-to-face conversation. It is presumed that we are all able to communicate fluently in normal face-to-face situations, and most people certainly do master this form of communication excellently. But it is not necessarily true that we are fluent communicators in all situations. When people must conduct a face-to-face conversation over a videotelephone, unexpected problems arise because the videotelephone actually changes the way we carry on a conversation, besides posing problems of mastering the technology (i.e. call set-up and adjusting terminal functions). These problems warrant a closer study.

In a live face-to-face conversation you will nod your head and use your hands and other forms of body language to show that you are listening and to indicate your approval or disapproval. In the sound telephone you have to resort to other strategies to keep the conversation flowing, such as 'acoustic nodding' or 'audio prompting', i.e. substituting sounds for visual signs to show that you are still on the line and paying attention.

When speaking on the videotelephone, we tend to continue using the acquired audio telephone habits, even though we should revert to our natural conversation strategies – the telephone habits are hard to get rid of and are perpetuated in new situations where they may actually be detrimental to communication.

Figure 4 The Tandberg videotelephone developed in co-operation with Norwegian Telecom Research (now Telenor R&D) in 1991

It took over a hundred years to learn to use the ordinary telephone and there is a continuing need for educating users as new technology and new services and functions are added. Adding to this the complication of transmitting and receiving a live picture will not make things easier for the user of the videotelephone.

When designing the user-interface for the videotelephone, some basic design goals were set up to guide us in our task: The videotelephone had to be easy to understand and easy to use, the procedures had to be easy to learn by first-time users and easy to remember by infrequent users.

2.1 Eye contact vs. face contact

The issue of eye contact in videotelephony has been hotly debated. It is generally believed that eye contact is just as important in videotelephone communication as in natural face-to-face conversation. Some ingenious designs have been proposed to overcome the problem of parallax error resulting from the need to mount the video camera outside the viewing screen.

However, recent findings do seem to indicate that *face* contact may be sufficient for obtaining proper communication between two people over the videotelephone [3]. In videotelephones with a low bit rate the diminutive screen size and the reduced screen resolution may preclude proper eye contact in any case, thus rendering any technical solution to avoid parallax unnecessary.

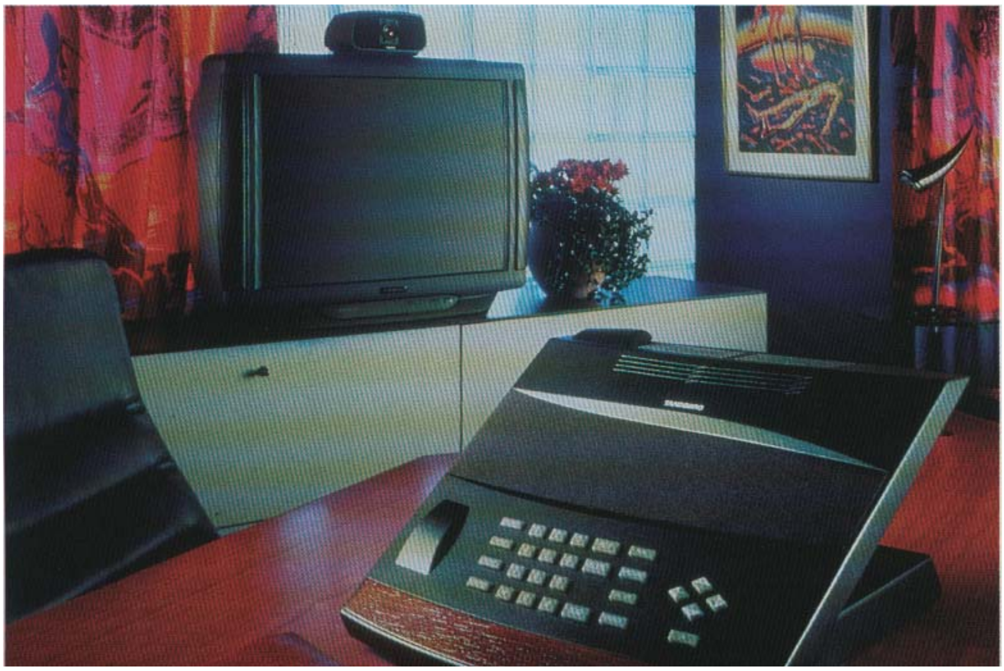


Figure 5 Another Tandberg state of the art videotelephone with detached display screen and camera. The Tandberg Vision 770

The important design rule is to keep the camera lens as close as possible to the edge of the screen. Most viewers prefer a picture where the camera looks slightly downward from above the screen, rather than looking slightly upwards from below the screen.

2.2 Compliance with ordinary audio telephony

The basic user-procedures in ISDN videotelephony should comply with those adopted for ISDN telephony and conform as closely as possible to the well established user-procedures in ordinary audio telephony. Any changes made to the traditional user-procedures should only be allowed where no other practical alternative is possible (e.g. en-bloc dialling, hands-free off-hook/on-hook, video enable/disable, etc.).

2.3 Number keypad

The layout of the videotelephone dialling keypad should conform to the standard layout specified in the *ITU-T Recommendation E.161*. This will ensure consistency with ordinary telephony (which will be in parallel use for many years) and it will avoid confusion and errors, especially for visually disabled people.

2.4 Ease of learning

Rapid learning of the end-user procedures for basic videotelephony is an essential design goal. The aim is to make a videotelephone call as similar as possible to an ordinary telephone call. The user-procedures in videotelephony should support the conventional face-to-face conversation between two people, without detracting their attention with complicated control procedures and status indications.

2.5 Left-to-right control layout

The layout of keys and controls should be organised in a left-to-right progression, i.e. keys and controls that are used first to initiate a call should be placed furthest to the left, and those that are used later, during a call, should be placed further to the right. It may not always be possible to do this since some controls are used both before and during a call (e.g. self-view).

2.6 Grouping according to function

Keys and controls should be logically grouped according to their function. The groups should be easy to discern and recognise, both visually and tactile. Groups should preferably have keys in contrasting colours or different sizes to make them more conspicuous.

2.7 Simplicity and consistency

Simplicity and consistency must be maintained. The number of keys and controls should be kept to the absolute minimum consistent with the basic functions of the videotelephone. Too many new keys, labelled with new function names or strange symbols, are certain to confuse most users. On the other hand, it is not advisable to use so few keys that each key must perform several functions.

2.8 Context independence

Keys and controls should not be context dependent, i.e. they should not have one function when the terminal is in one mode and a different function when the terminal is in another mode. Keys that change function according to context will confuse most people and increase mistakes. An exception to this rule are the number dialling keys, as they may be used both for dialling and for entering number codes and digits in other modes.

2.9 Initiating a call – 'off-hook'

Since videotelephones must also function as ordinary ISDN telephones, videotelephones must have a handset. On the ordinary telephone, a call is initiated by lifting the handset off-hook. The videotelephone handset must retain all the functions that it has in ordinary telephones, i.e. going into dialling mode when the handset is lifted off-hook and terminating the call when replaced on-hook.

When the videotelephone is in a *hands-free* or *loudspeaking* mode, it must be possible to initiate the terminal for a call, i.e. to go off-hook, without first having to lift the handset. Thus, a separate control or off-hook key will be required to initiate hands-free or loudspeaking videotelephone calls.

2.10 En-bloc dialling

In those ISDN services where *en-bloc* sending of the telephone number is used, an extra function control, i.e. a *delimiter* key for 'sending' the number, is required.

Experimental evidence from Norwegian Telecom [4] highlights the high design sensitivity of this solution and the importance of proper user-guidance in the dialogue. Various visual and acoustic prompts may be essential to inform the user that the delimiter (i.e. the 'send') key must be activated after a telephone number has been entered to actually transmit the number to the exchange. With the widespread use of GSM mobile telephones this concept of en-bloc dialling is fast becoming common knowledge and the use of the 'send' number/ 'off-hook' key and the hook-on key is fast gaining acceptance.

2.11 Self-view

A *self-view* function, i.e. to be able to see yourself on the screen, is mandatory in videotelephony. Self-view is necessary for positioning yourself correctly in front of the camera. When self-view is in use it should not stop transmission of the outgoing video picture, since this may disrupt communication.

If the picture is displayed the same way in self-view as it is transmitted the user will have problems because movements will be 'reversed' (if you move left, the image on the screen moves right, and vice versa). Ideally, the self-view picture should therefore be reversed left-to-right, like the image we see of ourselves in a mirror, to facilitate getting into the right position in front of the camera. When documents or objects are shown in self-view by the internal camera on the screen of the videotelephone, a mirror image will make it difficult to e.g. read printed material on the screen, but a mirror image will still make it easier to position the material properly before the camera.

The self view key should be a latching key (i.e. push once to activate, push once more to turn the function off) and it should have a visual indicator to show when the function is on.

Figure 6 Mock-up of a design exercise for a flat screen videotelephone from a student at the School of Industrial Design, Oslo

2.12 Visual reciprocity in face-to-face communication

In face-to-face videotelephone calls *visual reciprocity* should apply as the default mode, i.e. whenever A can see B, B should be able to see A. However, during a call objects and pictures may be transmitted from a subsidiary video camera by either party without inhibiting the picture received from the other party.

Also, when one party makes use of the self view function, the video picture in the other direction should still be transmitted normally, even though one party cannot see the other party when self-view is active. If full visual reciprocity must be adhered to at all times, it may seriously disrupt the natural flow of communication and make several intended uses of the videotelephone impractical or impossible.

Videotelephony is potentially more intrusive than audio telephony. To protect privacy, it is important that clear visual and/or acoustic indications are provided to advise both the caller and the called party of the video call. For the benefit of blind and visually disabled people, special acoustic indications are necessary.

2.13 Video picture enable/ disable

An important design goal is that there must be no possibility for either party to inadvertently send video and thus, unwillingly, being seen by the other party.

To ensure protection of privacy, either party must be able to stop outgoing video, while keeping the audio channel open, by using a special control key, either during call set-up or during the call. When a key is used it should have a latching action and a clear visual indication showing whether the function is active or not.

Figure 7 The use of a videotelephone is not always appropriate. Issues of privacy are important in videotelephone design

2.14 Terminating a call – 'on-hook'

It is important that a well defined and secure termination-of-call procedure is provided, both to avoid unnecessary expenses and to protect against intrusion on privacy.

In handset mode, the call is terminated by putting the handset back 'on-hook', as in ordinary telephony. In a loudspeaking hands-free mode, the call will have to be terminated by activating a special termination-of-call or 'on-hook' key.

Since it is much easier to forget to operate an 'on-hook' key than to put a handset back 'on-hook', a visual prompt may be needed to alert the user that the call has not been disconnected.

3 Videotelephony – the new communication form

Videotelephony is different in many ways from ordinary sound telephony. Ordinary telephony has been with us for more than a hundred years, but there are still people who feel uncomfortable when using the telephone; such skills have to be learned and developed.

Hopefully, it may not take another hundred years before we all will master videotelephony, but it is a new mode of interpersonal communication that has to be learned. As we get used to the videotelephone we will develop skills in using this new medium, and in due course we may become as proficient in its use as we are today with the sound telephone. All our experience so far has clearly shown that there are great potentials in video communication and that the procedures necessary to use a videotelephone can be quickly learned without too much trouble once its usefulness in various applications is appreciated.

References

1 Tetzner, K. Das erste Fernsehtelefon der Welt. *Funkschau,* 58, 42–44, 1986.

2 Norman, D A. *The psychology of everyday things.* New York, Basic Books, 1988.

3 Kellner, B et al. Bildtelefon mit Blickkontakt? – Technische Möglichkeiten und empirischen Untersuchungen. *NTZ,* 38, 698–703, 1985.

4 Hamnes, K, Nordby, K, Anderson, D. *Testbed for user-system interaction.* Paper presented at 13th International Symposium on Human Factors in Telecommunications, Turin, September 1990.

Anderson, D M, Hamnes, K. *A note on block sending. Reporting on results of pilot tests.* ETSI TC HF 1 (temporary document), 1990.

Colins, S C, Rabbitt, P M A. *Age and design.* Age & Cognitive Performance Research Centre, University of Manchester, 1990.

Galer, I (ed.). *Applied ergonomics handbook. Second edition.* London, Butterworths, 1987.

Maddix, F. *Human-computer interaction : theory and practice.* New York, Ellis Horwood, 1990.

Shneiderman, B. *Designing the user-interface. Strategies for effective human-computer interaction.* New York, Addison-Wesley, 1987.

Knut Nordby is Senior Research Psychologist at Telenor R&D, Kjeller, where he has worked since 1985. He is Chairman of the ETSI Technical Committee Human Factors (TC HF), of which he was a founding member. He is member of the Permanent Steering Committee of Human Factors in Telecommunications (HFT) and Adjunct Professor teaching Man-Machine Interaction at UNIK, Center for Technology at Kjeller, University of Oslo.

e-mail: knut.nordby@fou.telenor.no

A tactile marker for telephone cards: – From idea to global standard

KNUT NORDBY

1 Background

'Smart'-cards, i.e. machine readable ID-1 type cards bearing an IC (integrated circuit) chip or a magnetic stripe, are being used in an increasing number of societal applications, both in the professional and leisure spheres. Applications cover such diverse areas as finance (bank cards, credit cards, debit cards, 'electronic purse' cards), access control and security (admission cards, hotel key-cards, car-park cards, security lock cards), public transport (season tickets, taxi-cards, commuter passes), customer loyalty cards (petrol cards, frequent flyers cards, shopping benefit cards), medicine and social services ('medicards', social security cards, ID cards), leisure interests (library cards, ski slope or swimming pool day-pass cards, club membership cards) and, not least, telecommunications (prepaid telephone cards and SIM cards for GSM telephones) – the list can be extended indefinitely (see Figure 1). When used, most of these cards must be inserted in a specific orientation into the card-reading device or pulled in a particular manner through a swipe-card reader.

This ubiquitous 'plastic' card was first introduced in the 60s when a number of restaurants in New York agreed on a credit service for selected customers whereby these could take guests to lunch or dinner and be billed later. As proof of membership, each customer was issued a plastic membership card with embossed name and membership number. From this card a carbon copy impression could be made with a roller on the receipt to be signed by the cardholder. This system proved so successful that many more restaurants, hotels and various shops joined the system – the *credit card* had been born – and today hundreds of millions of cardholders can use their cards with millions and millions of vendors.

Figure 1 Machine readable cards have invaded a large number of professional and leisure activities in society, and new applications are added each day. However, many people experience problems in orientating and inserting ID-1 cards correctly in the card readers, irrespective if these are electronic door locks in hotels or automatic banking machines, often making several attempts before finding the right orientation

RUNE LISLERUD/SAMFOTO

Figure 2 The most common and frustrating problem for all users of mini-banks (automatic tellers to some) and similar technology, is to insert the card in the correct orientation. By putting a tactile marker on the card, it will be much easier to turn the card to the proper orientation before inserting it in the card-reader. This will not only make the service both simpler and friendlier to users, but may be essential to blind and visually impaired users

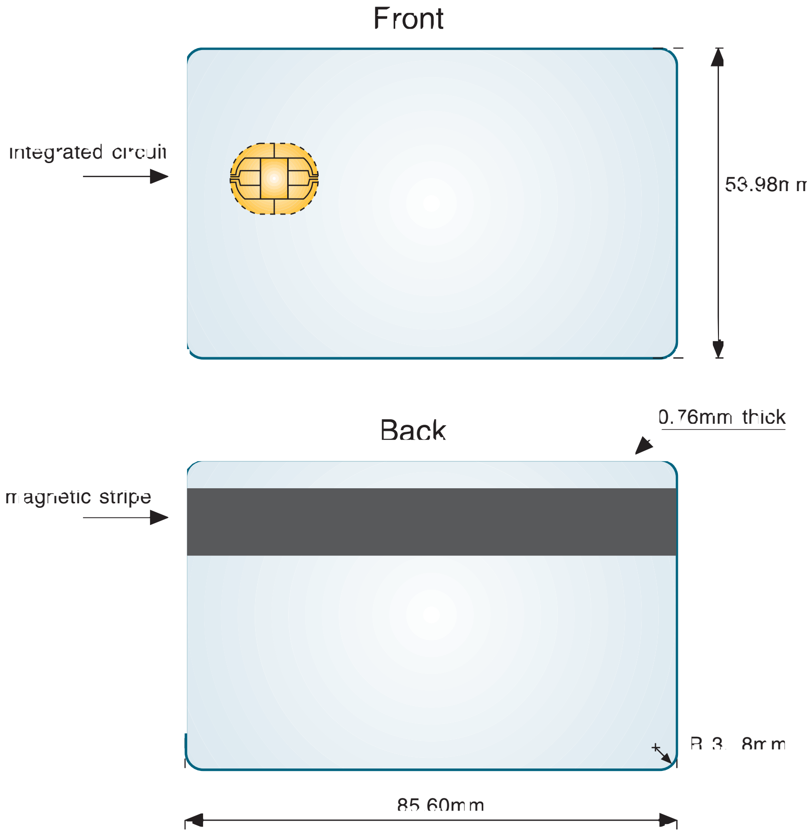


Figure 3 The measures of ID-1 machine-readable cards according to ISO standards 7810 and 7811

The physical dimensions of the card (see Figure 3) were inherited from the embossed metal printing plates used in address printing machines (e.g. *Address-O-Graph*) and have now been standardised (ISO, 1985) (see Figure 3).

By applying a magnetic stripe to the upper rear side of the card (above the embossing), cards were made *machine-readable*, and soon the card was taken up as a personal identification (ID) card by banks and in a large number of other security applications. When a programmable, electronic, integrated circuit chip was incorporated into the front left-hand side of the card, it became possible to read and write large amounts of information on the card, paving the way for the modern 'smart'-card, which has also been standardised (ISO, 1985).

This 'smart'-card technology was soon adopted by the telephone operators and is now used mainly for the common pre-paid telephone-card and the SIM (Subscriber Identification Module) card for GSM telephones. In the telephone card, a specific number of tariff units have been electronically stored on the card. When the card is used, these units are deducted one by one until the card expires.

The SIM holds information in the IC chip to authenticate a subscriber to use a particular GSM subscription. Initially, the SIM card was the same size as the ordinary ID-1 type telephone card, but, as mobile phones grew smaller and smaller, the SIM was too large to fit inside. A small piece of the SIM card, bearing the IC chip with contacts, can now be broken off the card and installed in the GSM telephone. This plug-in SIM module is a bit small to be handled comfortably by adult human hands, but normally this is not done very often.

As smart-cards are becoming more and more common, it is essential that the cards can be utilised by *all* cardholders. One important prerequisite for using an ID-1 card is that it can be orientated correctly for insertion into the card-reader. People who cannot orientate such cards correctly by visual cues alone, e.g. blind and visually impaired people, may thus be excluded from using important or even vital societal services (e.g. withdrawing cash, buying the daily necessities, using public transport, or making telephone calls).

To aid blind and visually impaired users to insert ID-1 type telephone cards correctly in the card-reader, the cards must be physically marked in such a way that their orientation can be determined unambiguously by the sense of touch.

Although such a *tactile marker* is essential for people who cannot rely on visual cues, people with no visual impairment may also experience problems with card orientation under adverse conditions, e.g. low illumination or inclement weather, and should benefit from a tactile marker to orientate cards. In situations involving safety (e.g. using a card while driving or while performing a hazardous operation) vision should *not* be used to orientate the card. Although the sense of touch is utilised in a large number of daily activities, e.g. identifying one key among several on a key-ring or finding the right coin among several in the pocket, tactile information is clearly an under-used information mode, compared to the visual or auditory modes.

There are at least six different tactile markers currently in use on ID-1 type telephone cards around the world (see Figure 4), but most telephone cards (like all other ID-1 type cards) do not have any tactile marker. The time was thus long overdue for creating an international standard for a common tactile marker on all ID-1 type telephone cards that must be presented in a specific orientation to the card reader. *European Telecommunication Standards Institute* (ETSI) Technical Committee Human Factors (TC HF) started work in 1994 with the aim of approving a common harmonised standard for a tactile marker for ID-1 type telephone cards. Thin flexible (TFC-1) cards, which have the same dimensions as ID-1 cards (e.g. like the telephone cards currently used in Italy or the tickets on the London Underground) are not covered by this standardisation, but the same general principles should also apply to these cards.

Liaison was established with the *European Committee for Standardisation* (CEN) TC224/WG6 to agree on a common standard for all ID-1 type cards to avoid different standards for telecommunication cards and banking cards. In 1992, work was also started in CCITT, Study Group 1, Question 18 (now *International Tele Union* (ITU-T) Study Group 2, Question 17), with the aim to adopt a global recommendation identical to the ETSI standard tactile marker for ID-1 type telephone cards.

Figure 4 Examples of various tactile markers on pre-paid telephone cards that are currently in use or have been in use in Europe. The broken off corner on Italian thin flexible (TFC-1) telephone cards is strictly not a tactile marker but is the result of authorising the use of the card to prevent fraud

2 User tests

As part of the standardisation work and to provide empirical data, user tests were carried out by Norwegian Telecom Research (now Telenor R&D) to determine: 1) Whether or not a tactile marker actually will aid users to orientate telephone cards correctly, and, 2) which of seven tested tactile markers was best for this purpose. In these tests the subjects' task was to sort stacks of ID-1 cards, with and without tactile markers, to the same orientation.

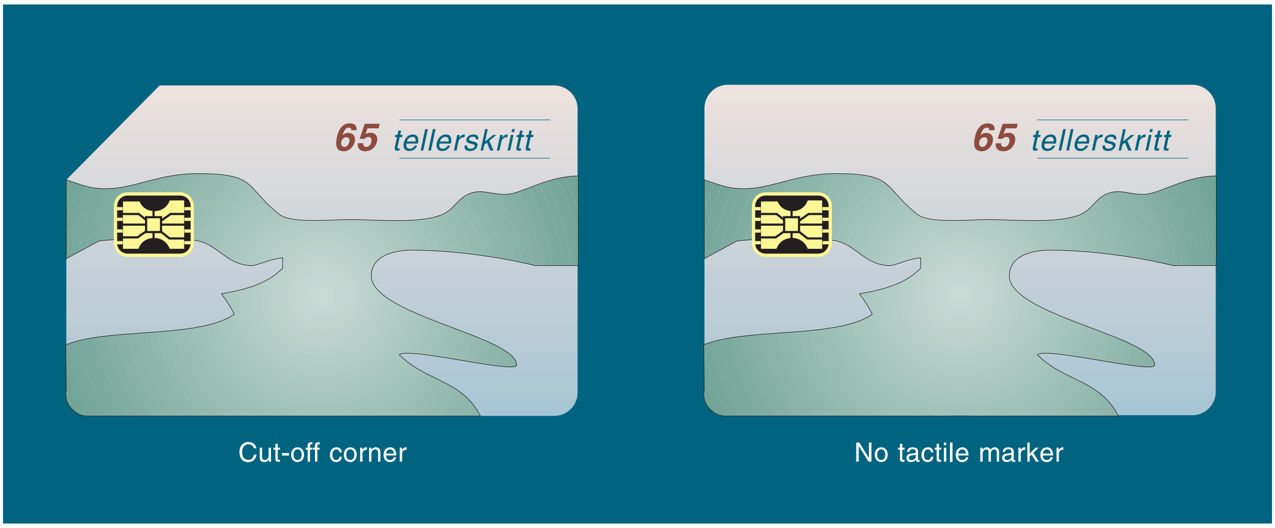


Figure 5 The two ID-1 card types used in Test One

A total of 92 subjects participated in the two tests. They included people who were blind (N = 17), partially sighted (N = 33), elderly, mentally retarded and people with no apparent disability (N = 42), among them some 16–17 year old students. Some of the subjects had and some had not used ID-1 cards before, but none of the subjects had previously used cards with tactile markers.

2.1 Test One

In this test the subjects (N = 75) were presented with two stacks of 20 randomly oriented, graphically marked, telephone cards. The cards in the *experimental stack* had a tactile marker in the form of a *cut-off* corner (see Figure 5, left panel), while the cards in the *control stack* had no tactile marker (see Figure 5, right panel).

The subjects' task was to sort and re-stack the cards in each stack so that they were all oriented in the same direction. The task was to be completed as fast as possible. When re-stacking the cards with tactile markers, the subjects were instructed to "stack the cards in the same orientation so that the tactile markers were on top of each other". When re-stacking the cards with no tactile marker, the subjects were instructed to "place the cards 'picture-side-up' with the text on each card oriented in the same direction". The subjects were also asked to give a subjective *preference ranking* of the two card types (1 for best and 2 for worst).

The time required to perform the task, the number of errors made in each task and the card preference rankings were recorded. Mean time used, mean number of errors made and mean preferences ranking for each card category were computed. The results from *Test One* are shown in Table 1. Paired samples T-tests between the mean time used to complete the task ($t = 6.60$, $p < 0.01$), the mean number of errors made ($t = 5.79$, $p < 0.01$) and the mean preference ranking (1.06) for the card with the cut-off corner clearly show that a tactile marker *significantly* aids all users in orientating ID-1 cards quicker and with fewer mistakes.

Of the 17 gravely visually impaired subjects who participated in the test, 15 could not complete the task of re-stacking the visually marked cards in the control stack, since they obviously could not make use of the visual markings on the cards. Their data are not included in the results. However, they could all easily execute the task of re-stacking the cards with a tactile marker in the experimental stack. For blind users, a tactile marker is *essential* for orientating ID-1 type cards.

Table 1 Paired samples T-test, Test One (N = 75)

UNMARKED$_{ERROR}$ VS. MARKED$_{ERROR}$	UNMARKED$_{TIME}$ VS. MARKED$_{TIME}$
H_0 = UNMARKED$_{ERROR}$ = MARKED$_{ERROR}$	H_0 = UNMARKED$_{TIME}$ = MARKED$_{TIME}$
H_1 = UNMARKED$_{ERROR}$ > MARKED$_{ERROR}$	H_1 = UNMARKED$_{TIME}$ > MARKED$_{TIME}$
MEAN DIFFERENCE = 1.667	MEAN DIFFERENCE = 19.107 s.
SD DIFFERENCE = 2.495	*SD* DIFFERENCE = 25.059
$T = 5.785$, $DF = 74$, $p < 0.01$	$T = 6.603$, $DF = 74$, $p < 0.01$

2.2 Test Two

In *Test Two* the subjects (N = 92) were presented with seven different stacks of 20 randomly oriented, white ID-1 type cards. Each of the stacks had a different tactile marker: 1) *Small diameter notch;* 2) *2 mm notch;* 3) *Embossed arrow (Braille dots);* 4) *0.7 mm notch;* 5) *Hole-in-card;* 6) *Cut-off corner,* and 7) *0.3 mm notch* (see Figure 6 for details).

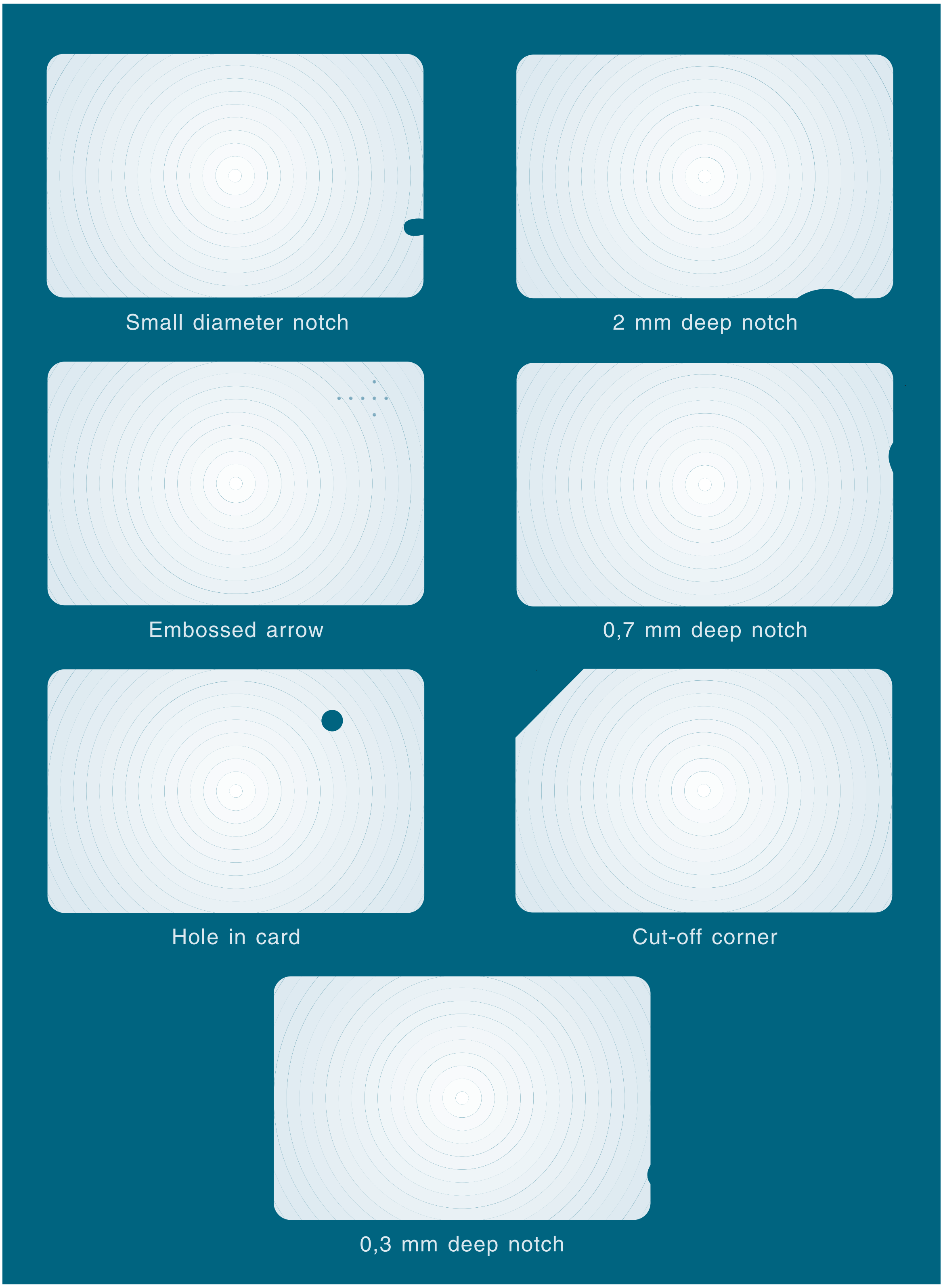


Figure 6 The seven tactile markers used in Test Two

Table 2 *The results from Test Two (N = 92)*

	0.7 mm notch	Cut-off corner	0.3 mm notch	2.0 mm notch	Hole in card	Emboss arrow	Small notch
N	77	92	92	91	92	92	89
Mean error	0.72	0.17	1.62	0.08	0.06	0.13	0.55
N	79	92	92	92	92	92	91
Mean time	64.86	44.41	66.33	47.15	47.55	55.26	49.36

The subjects' task was to orientate and re-stack the cards in each stack so that they were all oriented in the same direction. The task was to be completed as fast as possible. When re-stacking the cards, the subjects were instructed to "stack the cards in the same orientation so that the tactile markers were on top of each other". The subjects were also asked to give a subjective preference ranking of each tactile marker (1 for best and down to 7 for worst).

The time to perform the task, the number of errors made and the tactile marker preference ranking were recorded for each card stack. Mean time used, mean number of errors made and mean preference ranking for each tactile marker was computed. To minimise any position or learning effects, *Test One* and *Test Two* were randomly assigned, and within tests the presentation order of the different card stacks was randomised.

The results from *Test Two* show clear differences between the seven tactile markers (see Tables 2 and 3). It should be noted that the mean time required to orientate the cards with the *cut-off corner* in *Test Two* is nearly identical to the time obtained for cards with this tactile marker in *Test One*, thus confirming the inter-test reliability. The results obtained with the *0.7 mm notch* and the *0.3 mm notch* are similar to or worse than the results obtained with the *unmarked* cards in *Test One*. Thus, no real benefit would be gained by using these two tactile markers, except for young blind people with a good sense of touch who are totally dependent on *any* form of tactile marker. It should be noted that most blind people are elderly who have reduced skin sensitivity, and most people who are blind in the western industrialised societies (Europe and North America) become blind as the result of untreated *Diabetes mellitus*, which often results in loss of vision through *retinopathy*, a vascular disorder of the retina. But diabetes also destroys the capillaries of the skin, thus reducing the sense of touch. Although some tactile markers came better than others in *Test Two*, not all can be recommended for various technical reasons.

Table 3 *Time, error and rankings in Test Two*

Overall ranking	All subjects (N = 92)			Blind subjects (N = 50)		
Tactile marker	**Time**	**Errors**	**Rank**	**Time**	**Errors**	**Rank**
1. Cut-off corner	1	4	1	1	4	1
2. 2.0 mm notch	2	2	2	3	2	2
3. Hole in card	3	1	3	2	3	3
4. Small dia. notch	4	5	4	4	6	5
5. Embossed arrow	5	3	5	5	1	4
6. 0.7 mm notch	6	6	7	6	5	7
7. 0.3 mm notch	7	7	6	7	7	6

2.2.1 Cut-off corner

The *cut-off corner* card required the least time to orientate. In terms of errors, it came only fourth, but it should be noted that one subject was responsible for 11 of a total of 16 errors made and that the differences in number of errors between the four best markers is small compared to the number of errors made in the worst cases. The *cut-off corner* card was ranked first by most subjects.

Overall, the *cut-off corner* was judged best in *Test Two*. For technical reasons, however, it cannot be used on ID-1 telephone cards (it corrupts the magnetic stripe and makes card feeding difficult). It is currently in use on thin flexible (TFC-1 type) telephone cards in Italy.

2.2.2 2 mm notch

The *2 mm notch* was second in terms of time used, second in terms of errors made and second in the preferences ranking. Overall, the *2 mm notch* was judged second in *Test Two*. There are no valid technical objections against its use on ID-1 type telephone cards. This tactile marker is currently in use in Denmark.

2.2.3 Hole in card

Cards with a hole came third in terms of time used, first in terms of errors made and third in the user preference ranking. Overall, *the hole in card* was judged to be third in *Test Two*. For technical reasons, however, this tactile marker cannot be recommended since the hole corrupts the magnetic stripe area and any future optical storage area.

2.2.4 Small diameter notch

The *small diameter notch* was fourth in terms of time used, fifth in terms of errors made and fourth in the user preference ranking. Overall, the *small diameter notch* was judged to be fourth in *Test Two*. There are no technical objections against its use on ID-1 cards. This tactile marker was used on telephone cards in Norway during a short trial period.

2.2.5 Embossed arrow

Cards with an *embossed arrow* in raised dots came fifth in terms of time used, third in terms of errors made and in the user preference ranking. Overall, the *embossed arrow* was judged to have come fifth in *Test Two*. However, this

PER-ANDERS ROSENKVIST/SAMFOTO

Figure 7 The prepaid telephone card is a most useful service both to the operators, who do not have to collect the coins from the terminals or have them vandalised by people who are after the cash, and to the users who do not have to worry about carrying the right coins. However, it is important that the users manage to insert the cards in the right orientation. Applying a tactile marker to cards simplifies insertion and makes it possible for blind and visually impaired people to utilise the service

Table 4 Overall rankings in Test Two

All subjects (N = 92)	Blind & visually impaired subjects (N = 50)
1. Cut-off corner	**1.** Cut-off corner
2. 2.0 mm notch	**2.** 2.0 mm notch
3. Hole in card	**3.** Hole in card
4. Small diameter notch	**4.** Embossed arrow
5. Embossed arrow	**5.** Small diameter notch
6. 0.7 mm notch	**6.** 0.7 mm notch
7. 0.3 mm notch	**7.** 0.3 mm notch

tactile marker cannot be used as the increased thickness of the card prevents its insertion into card-readers and the magnetic stripe area and any future optical storage area are corrupted.

2.2.6 0.7 mm notch

The *0.7 mm notch* was sixth in terms of time used, sixth in terms of errors made and seventh in the preferences rating. Overall, the *0.7 mm notch* was judged sixth in *Test Two*. This tactile marker was proposed by CEN TC 224/WG6. It should be noted that this marker did not do any better than the unmarked cards in *Test One.*

2.2.7 0.3 mm notch

The *0.3 mm notch* was seventh both in terms of time used and errors made and sixth in the preferences rating. Overall, the *0.3 mm notch* was judged seventh and last in *Test Two*. This tactile marker has been used on several telephone cards in Europe. This marker did not do any better than the unmarked cards in *Test One*.

2.3 Blind and visually impaired subjects

Of the 92 subjects who participated in the trials, 17 were completely blind and 33 were visually impaired. The results obtained with these subjects (N = 50) were singled out for separate analysis. The results are practically the same as for the whole group. Time and error rankings for both groups are shown in Table 3 and subjective ratings in Table 4.

An interesting observation, which we have not seen reported before, was that blind people, when exploring the cards with their fingers, felt the edges (usually the long edges before the short) before they felt the surfaces of the cards. Thus, a tactile marker on the edge of the card should be detected before a marker on the surface of the card. This also applies to people with normal or impaired vision when handling a card prior to insertion in a card-reader.

Figure 8 The first two telephone cards, from British Telecom and Telenor, to bear a tactile marker according to ETSI ES 300 767 and ITU-T Recommendation E.136

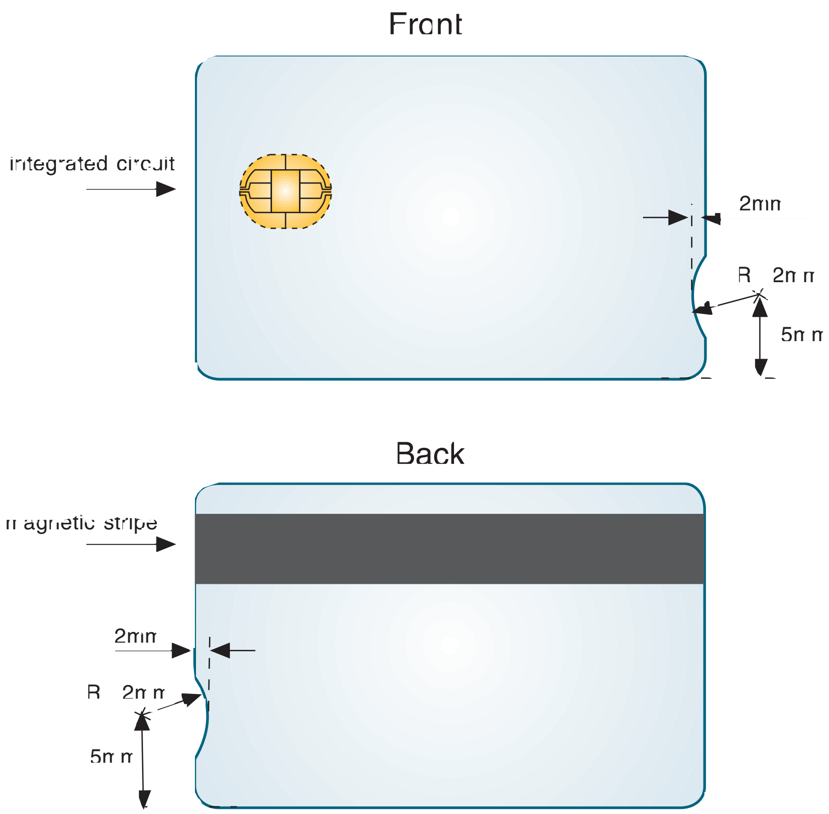


Figure 9 The measurements of the tactile marker for telephone and SIM cards according to ETSI ES 300 767 and ITU-T Recommendation E.136

2.4 Recommendations

The results from both tests showed that a tactile marker on ID-1 type cards was preferred to not having a tactile marker, and a tactile marker significantly aids card orientation for *all* users. The results from *Test Two* clearly show that the more prominent or salient the tactile marker is, the better it is. Although the *cut-off corner* marker came best in the trials, slightly ahead of the *2 mm notch*, it could not be recommended for technical reasons. Thus, the *2 mm notch* was the clear recommendation, both on usability and technical grounds. The results from the full group and the results from the sub-group of blind and visually impaired subjects are virtually the same (see Tables 3 and 4) and our recommendation is the same for both groups. The two tactile markers that came last in the trials (*0.7 mm notch* and *0.3 mm notch*) were no better than unmarked cards.

3 The standards

On the strength of the results of the user trials, the *2 mm notch* was proposed in the first ETSI Draft Standard as a European regional standard for a tactile marker for telephone and SIM cards. Also ITU-T proposed this tactile marker in the first draft recommendation. However, the liaison with CEN TC224/WG6 made it clear that the chosen position on the lower long edge was not acceptable since it would interfere with the transport of cards in some automatic banking machines. Also, CEN wanted a much shallower notch than 2 mm (initially they opted for 0.3 mm, but eventually accepted 0.7 mm). Our trials had shown that the 0.3 and 0.7 mm notches were no better than unmarked cards (with the exception of young blind people without any other impairment). The position of the notch, however, was open to discussion. A test was carried out by Mr. T. Cloke of BT Labs at Martelsham Heath, UK, to determine which of the two positions of the *2 mm notch* was more acceptable. The two positions were the original *2 mm notch* position on the lower long edge of the card and the *0.3 mm notch* position on the right hand short end of the card. The test showed that the position on the long bottom edge was preferred to the position on the short right hand end, but the diffrence between the two positions was small and not significant.

On the strength of this finding ETSI changed the position of the tactile marker in the next version of the Draft Standard, and was soon followed by ITU-T 17/2. CEN TC224/WG6, however, came up with a proposal of a 0.7 mm deep notch *on the long bottom edge* of the card, i.e. the same position as the original 2 mm notch! This led to some confusion. However, ETSI TC HF and ITU-T 17/2 could not hold up their standardisation activities any longer and the proposals for a 2 mm deep notch on the right hand short end of the card went through the formal adoption procedures (see Figure 9).

On Monday 26 May 1997, the ITU-T Study Group 2 plenary meeting in Geneva unanimously adopted IUT-T Recommendation E.136 “Tactile Identifier on Pre-paid Telephone Cards” which thus became a global standard – thereby beating ETSI by a month.

On Friday 27 June 1997, ETSI Vote 97/26 ended with the result that the prETS “Human Factors; Telephone prepayment cards; Tactile identifier” was adopted as ES 300 767. As this goes to press, BT has decided to adopt the ETSI/ITU standard tactile marker for their telephone cards. Other operators are sure to follow suit (Figure 8).

In a belated proposal from CEN TC 224/WG6 on prEN 1332-2 “Identification Card Systems – Man-Machine Interface Part 2: Definition of a Tactile Identifier for ID-1 cards”, it was now proposed to cut off the lower right corner radius (which previously had been ruled out as totally unacceptable) as an alternative to the 2 mm deep edge indentation. Whatever kind of tactile marker is adopted, some party may find objections.

The main objection against a tactile marker has been cost. The cost of applying a tactile marker is negligible compared to the cost of the whole card. The

Figure 10 Promotion telephone card issued for the 16th International Symposium on Human Factors in Telecommunications, Oslo, Norway, 12–16 May, 1997 (HFT'97)

main objection against the new standards is that some card manufacturing equipment for transporting the cards from station to station may have to be modified, but compared to the expected growth in revenue from increased card usage these costs should be minimal, especially since such equipment must be replaced regularly as it is worn out.

The important lesson here is that *user trials* are essential to determine the optimum form and location of a tactile marker on telephone cards to make the cards easier to orientate for all users. Although this may be in conflict with the interests of some industrial parties, the interests of the users should be given priority – in the end it is the customers who will determine if the telephone cards will be used or not.

References

ISO 7810. *Identification cards : Physical characteristics.* 1985.

ISO 7811/1. *Identification cards Part 1: Embossing.* 1985.

ISO 7811/2. *Identification cards Part 2: Magnetic stripe.* 1985.

ISO 7811/3. *Identification cards Part 3: Location of embossed characters oncards.* 1985.

ISO 7811/4. *Identification cards Part 4: Location of read-only magnetic tracks.* 1985.

ISO 7811/5. *Identification cards Part 5: Location of read-write magnetic track.* 1985.

ISO 7816-1. *Identification cards, Integrated circuit(s) cards with contacts Part 1: Physical characteristics.* 1987.

ISO 7816-2. *Identification cards, Integrated circuit(s) cards with contacts Part 2: Dimensions and location of contacts.* 1988.

ISO DIS 11694. *Identification cards, Optical memory cards Part 2: Dimensions and location of the accessible optical area.* 1995.

ETSI ETR 165. *Human Factors; Recommendation for a Tactile Identifier on Machine Readable Cards For Telecommunication Terminals.* Januar, 1995.

ETSI ES 300 767. *Human Factors; Telephone Prepayment Cards; Tactile Identifier.* 1997.

Knut Nordby is Senior Research Psychologist at Telenor R&D, Kjeller, where he has worked since 1985. He is Chairman of the ETSI Technical Committee Human Factors (TC HF), of which he was a founding member. He is member of the Permanent Steering Committee of Human Factors in Telecommunications (HFT) and Adjunct Professor teaching Man-Machine Interaction at UNIK, Center for Technology at Kjeller, University of Oslo.

e-mail: knut.nordby@fou.telenor.no

The User-Interface Design process; some new year resolutions

SIMON CLATWORTHY

This article looks at the interface design process and reflects upon its practice in projects. Methods and disciplines are discussed based upon recent project experience, and some future directions for interface design are suggested. The article focuses on systems aimed at the consumer market, although the distinction between home users and office users is now becoming less and less distinct. The paper takes a pragmatic and personal view of the design process, rather than an idealised view. This view is of course my own, and does not necessarily reflect the views of the organisations or projects I work for.

Introduction

The job of designing a good user interface just gets harder and harder. The rate of change in user interface development is increasing rapidly, even though the modern desktop is not so modern any-more (it must be reaching its 25th birthday soon). The reason that the design process has become more complex is a combination of factors such as increased technological complexity, new design disciplines, an increased focus on good interface design, and better user interfaces to compete against. It seems that finally, the Human Factors profession is in the spotlight with expectations greater than ever before that we can produce miracles. If Andy Warhol was right (everyone will be famous for 15 minutes), our 15 minutes of fame has started ticking, and I am afraid that the first 10 minutes passed without us realising it. We should therefore think carefully about how we use the remaining minutes.

The design process as she is practised

The design process is becoming more and more complex. Time to market is being squeezed at one end, whilst at the other end, an increasing number of experts need to be involved in the process and professionally co-ordinated. In addition, expectations of quality, novelty and 'sexiness' in interface design are increasing.

This Internet page from Mercedes shows a good combination of design and user interface skills. The design builds well upon Mercedes market profile, yet is very easy to understand and navigate through. Although there are many menu alternatives on this page, the screen does not seem cluttered due to grouping of functions and typography. A good example of a second generation web-page

However, our understanding of the design process is becoming clearer, particularly upon the focus of each of the different project stages and how each can be influenced. If I take a bird's eye view of the design process, then it has essentially three major phases which are critical to the user interface design: a user focus, a technological focus, and back again.

The following view of the design process is based upon recent project experience where human factors has had a reasonably large role. I still notice, however, that many projects have a tendency to skip the first stage and jump directly to technical development.

The first phase: user focused

The first phase of a project has very much a user focus. It defines goals, user groups/market segments, carries out analyses of needs, talks to users and generally creates a user requirement spec in one form or another. It also often ends up with some form of interface visualisation; a rapid prototype. At this stage the project is very much based upon a user model of a product or application with very little focus on implementation aspects. I call this phase the 'white knight' phase, it would like to be focused upon idealistic solutions and based purely upon user needs. Technical expertise has a back seat during this phase, increasingly too much so ("Yes we are sure it is important, but it comes later – we must not let it disturb the identification of *real* needs"). This phase usually takes quite some time and for good reason; accurately identifying user requirements and defining project goals takes time. We should respect this, but we somehow try to apologise and rush things through. Building a good project team takes time and cannot be rushed.

The second phase: technical implementation

The second phase of a project is the implementation phase, and has a clearly (often exclusively) technical focus. At this stage, the users, the marketing people and the interface designers usually have a back seat, and development moves quickly and seemingly efficiently. This stage of the design process is what I call the 'black box' phase. It takes user focused input, translates it and breaks it down into technical sub-components and develops solutions in parallel. The parts are then put together and tested to form a whole. The 'black box' phase of system development is resistant to user influence, and sometimes directly antagonistic towards it. At this stage, the project is very much based upon a technical model of a product or application.

The final phase: user focused

The final stage of a project is the completion, marketing and launch. This phase returns to a user focus, and many of the user disciplines from the first phase come back into the project. It is not unusual that users and experts alike receive a few surprises at this stage and find that the end result deviates from the intended result in a few major ways, and many minor ways. This stage I call the 'amphetamine' phase. Tight deadlines, technical problems and a product slightly wide of the market require everyone to work around the clock, on borrowed time. The final stage is not always a nice period in a project, which probably explains people's unwillingness to document it properly, and to carry out the all-important de-briefing afterwards.

This is very much a simplification and a generalisation, but it highlights one thing which I think we are not very good at; managing the shift of emphasis between a user focus and technical focus in projects. The fact that these stages are so clearly visible in projects is actually in itself a sign that something is wrong. What should we do about it? Well, here are my new year's resolutions for next year.

Resolution #1: Get more out of the early stages

I have noticed during the last year an increasing interest for a user focus at the early stages of a project, and an growing willingness to devote time to this area. However, projects increasingly lack a means of effectively documenting this, and even more importantly, extracting the relevant information from this phase. It is not unusual to see a user based task description in a project, yet find that nobody really has *analysed* this and said: "So what are the *consequences* for the project and for the user interface?". User requirement experts need to become much better in pulling out the implications of our analyses and communicating them effectively.

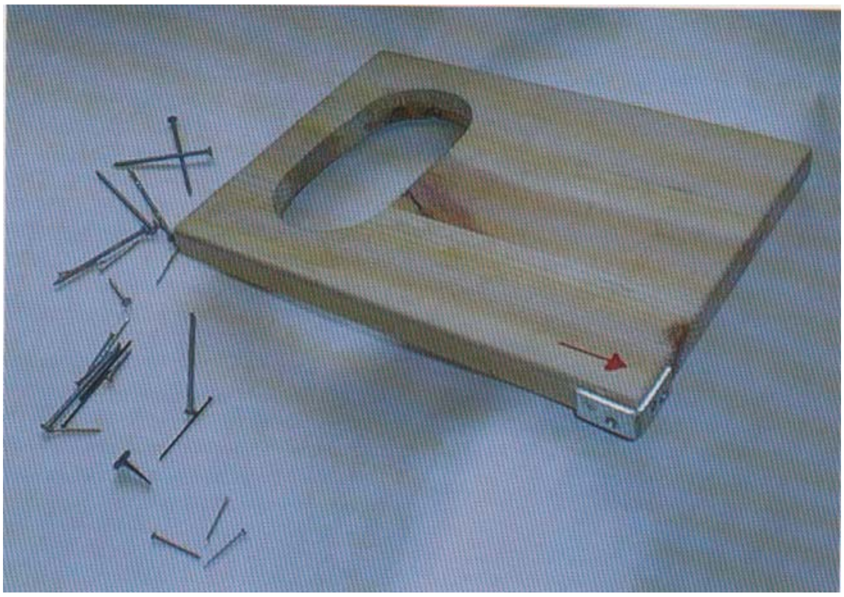
I really like the HUFIT tools (slightly reworked as USERFIT) in this respect. If you can get over the masses of forms which HUFIT requires, or even better, tailor your own, then there is a lot to be said for their approach of forcing a project to focus upon the discussion of the *implications* of user requirements in a project. Forcing yourself to think about these things together with the stake holders in a project (including the implementation team) and then incorporating them in the project mandate are some of the great things HUFIT has taught me (thanks, HUSAT). Thanks also to Ken Eason and Susan Harker for their simple yet effective task *analysis* method 'an open systems approach to task analysis'.

Resolution #2: Avoid virtual ergonomic teapots

This is a trap I have fallen into loads of times, and which I still have a tendency to fall into. It is based on the (mistaken) belief that poor user interfaces are the results of 'nerds' doing the designing, and that if only we (the user interface profession) could manage (read: control) the process, then everything would be perfect. It is just not true, although I wish it were.

During the 1960s, the human factors profession set about the design of *the* ergonomic teapot. Teapots are not particularly user friendly. The point of balance is wrong, you can scald yourself, they

The user interface design process is still too much based upon the relay race metaphor. We should start playing football instead



Gautam Ghosh from ISI as in Norway has a nice exercise in user requirement specification; the specification of a hammer, and the implementation of the spec. Well, you did say a metal tipped object which can be held in the hand and used to impart force to nails so that they can be driven into softer materials in a chosen direction ...

drip, you can never find the tea bag, etc., etc. The project therefore redesigned the teapot from a purely ergonomics perspective and resolved most of these user problems. They did however also produce one of the ugliest things in the world. Not only that, it was difficult to manufacture.

User interface designers, particularly the human factors discipline (and I am one of them), seem to want to create two kinds of 'virtual ergonomic teapots':

1. Interfaces which work well on paper, but which when technically implemented cause all sorts of problems and alienate users ('thou shalt have a good interface, but your system will suffer as a consequence').
2. Interfaces that are easy to use but which are incredibly boring ('thou shalt be efficient but not have fun').

What can we do about the problem? Well, I have a feeling that it might go away on its own after a while. The human factors/user interface profession lacks self esteem; it has always battled to be accepted as an important aspect of a product, and is just dying to have the opportunity to show what it really can do. Now it seems that we are increasingly being offered the opportunity to show what we can do, and not surprisingly are making a few mistakes as we learn. Hopefully, increased experience should temper our ambition.

Resolution #3: Play more football and run fewer relays

One of the problems with the three phases mentioned earlier is the transition stages from user focus to technical focus and back again. If you think of the design process as a relay race, it seems that the baton does not get handed over properly at the early stages, and that again, it does not get handed back properly towards the end.

In fact, the problem might be the idea that there is a hand over at all and the process should not be likened to a relay, but to a football match, more of a team effort. This may be the direction for future development, but at the moment we are still in the world of the relay race.

We should, however, be aware of the consequences of moving towards a football team approach. In a relay, only one person is working at a time, although that person is the best in the field. You can also replace one or all of the relay team up until or even during the race. In football, at times all eleven players are working together, and it is wise that all eleven have played together before the match starts. Additionally, substituting more than one or two players underway has a negative effect upon the result as a whole. Football is therefore much more costly, requires a team captain, and in addition a greater degree of strategy and co-ordination. There again, football players (at least the good ones) get paid a lot more, and receive much more media attention, so no prizes for guessing what kind of sports metaphor I am putting my effort into in future.

Resolution #4: Communicating that 'a spade is a spade' (or alternatively, a tool for soil excavation and transportation)

This resolution is just a recognition of the problem of transition from a user model to a technical model of an application. If we stick with the relay race analogy for the time being, the user centred 'baton' has a different shape from the technically focused 'baton'. The user focused 'baton' is based upon how the human mind works and how people best can carry out tasks. The technical 'baton' is based upon how computers store, retrieve and present information and how to structure bits to make the best use of this. Until processors and databases work in the same way as the human mind, there is going to have to be a transition from one to the other, and back again. As with all square pegs which have to go into round holes, bits go missing underway. The important thing is to try and make sure that the bits which get shaved off are not of critical importance. Many systems look good at the specification stage, but become a mess when implemented, mainly because the wrong bits went missing underway. This is a communication problem in the creation of a requirements specification, and we (user interface experts) need to be much better at clearly communicating the critical parts of a user requirements spec for projects, those which should be cast in concrete and the less critical, those which can be adapted.

We can learn a lot from the oil and international standards world here. They have a very clear notation for writing requirements, based upon the two magical words *shall* and *should.* By being very clear about requirements which shall be implemented as opposed to those which should be implemented, and receiving project approval for this, the move from a user focus to a technical focus can be improved.

Resolution #5: Stop lugging the piano and start a band

A famous jazz musician, who at one time played several instruments, but finally focused upon the saxophone, once said that "the best thing about just playing the saxophone is that you don't have to lug a piano around anymore". I get the feeling that user interface designers have been playing loads of instruments and have not only lugged around the piano, but also the double bass, saxophone and drums. More and more specialisms are needed to create good solutions, and being a sole representative of user interface design in a project is backbreaking. It is about time that we took the consequences of this and involved other experts in the process. I am happy to play just the saxophone in a good band (as long as I can play a few solos).

Resolutions #6: Combine DIY and experts for user requirement capture?

One of the questions always asked in a project is how much one should hire in experts during the project and how much the project team can do themselves (DIY).

Experts in requirements capture will be quicker, more thorough and will identify more requirements than non-experts. They will also be better at separating important requirements from less important (see #4, above). They do, however, have an overhead. They cost money (not always more than if you do it yourself, depending on hourly rates) and have to communicate the results to the development team.

DIY can lead to ineffective resource usage, the use of incorrect methods or the incorrect interpretation of results. However, the results become part of the implementation process, continuity is guaranteed, and the development process has an increased user focus as a whole.

Ideally, we should hire in experts in user requirements to form part of the project team and then keep them in the team throughout the design process. They should be responsible for managing user requirements capture, yet use project members to carry out some of the data collection. This requires the ability to manage their participation without wasting resources. If experts become too peripheral, then they may lose touch with project development (therefore weakening your football team). If you involve them too much, you may waste resources.

In a recent project I tried to improve communication with the implementation team during the requirements capture stage of a project by including them in the requirements capture (they carried out user interviews, wrote parts of the URS), and we held HUFIT workshops together to discuss results and critical aspects of the user interface design. This was working well, and we did at one stage create a common understanding of user and technical needs. However, I also discovered the vulnerability of such a team to change. At one stage of the project, several members of the implementation team were changed, and all of the careful work in creating a cohesive team went down the pan. Not daunted by this, I have vowed to continue this approach, but will now improve team building and cohesiveness. So, again, knowing how to do things correctly has become simpler; being able to carry them out, more complex.

Resolution #7: Widen the net

During the early days of user interface design, the focus was very much based upon the layout of menus and navigation within hierarchical systems. At this point we jumped up and down and said "computing is all about tasks and improving task performance". The net was duly widened to include this, and task analysis became a focal point. It gradually dawned upon us that no matter how well we focused on the task, it was not enough; the organisational context and environmental context were important. We jumped up and down about this and the net was ceremoniously spread wider to include this. Now, as we move towards computing at home, we are becoming aware of the need for a social and cultural foundation for user requirements. I am sort of half jumping at the moment about this and resolve to jump more next year to spread the net even wider. So far, we have not succeeded in including this factor adequately for the computing profession to take it on board, but it is coming.

I was promoting the use of HUFIT earlier, and believe that it has a lot to offer the design process. In some ways, the methods developed in MUSIC (original name: Measuring Usability In Context) for evaluation have taken some of the HUFIT philosophy further. In fact, one of the most useful parts of MUSIC, the context analysis, is something I have used during user requirements capture and it is a great way of structuring work. Now I am stumped, because I need to have some methods which can be used to step further from MUSIC's context analysis to include a social and cultural focus. The problem is, I cannot find one method which is design-process friendly. Tips, anyone?

Resolution #8: Move from the information society to the experience society

Peter Cochrane from BT has a point, we seem to be moving from the information society to the experience society. His slogan "instant gratification anywhere, anytime" does seem to be a little over the top

Table 1 The table shows how, over time, the focus of user interface design has moved from supporting tasks to supporting social structures. Each time we widen the net, we invoke complexity

Period	Theme	Challenge
1980–1990	Menu-based systems	Menu structures, Time delays, Navigation
1990–1995	Windows and GUI	Supporting tasks, organisational context
1995–1997	Internet	Social and cultural context
1997+	Internet at home	Family structures, computing as an experience

but he does have a point and it is going to have a huge effect upon user interface design.

> "With the gap between competing products narrowing in terms of performance and quality, the experience that a product delivers is rapidly becoming the key to offering distinctive value to the customer."
> *Harvard Business Review*

What does this mean for interface design? Well, it means that users will be expecting more from their interface than they do at the moment. The computer games industry has always had a focus on this aspect, so we can expect that some of the elements of computer games will be coming to an interface near you, and soon. In fact, Nolan Bushnell, founding guru of Atari®, is a strong voice in this respect. His article, *Relationships between fun and the computer business*, makes interesting reading.

It also means the introduction of new disciplines to the design process, particularly those from the marketing and media world. Internet has shown us the importance of capturing a user's attention within the first five seconds. Like it or not, this trend is going to accelerate, so jump aboard and put your seat belts on. In fact, the traditional human factors input to a project will probably lose some of its relevance and take a back seat to design and media disciplines as the goal for computing moves from an efficiency directed development towards an experience directed one.

Resolution #9: Prepare my employer for the consequences of resolution #8

I have a real dilemma here: how analytical should the design process for user interfaces be? Traditional Human Factors is analytical, whereas Design is a creative, almost artistic process. If I ask if the two can be combined then I always receive a hearty 'yes' as a reply. When I ask in which proportions they should be combined, and which one should come first, I do not receive many answers, and I am unsure myself. Obviously, the answer will vary from project to project, but then who should decide that project A should be 30 % Human Factors and 70 % Design, whilst project B has entirely different proportions?

I am also unsure of how much user requirements knowledge you should transfer to designers and how much total freedom should you let them have. One of the goals of using designers is both to pull in people from outside a problem area, and at the same time bring in creativity at a critical part of the design process. This is usually at the same stage as user requirements capture, and I have tried many variants over the years. Designers function well with some basic user requirements, but they become rapidly constrained when inundated with detail. Work with designers is often an evolutionary process, so you do not have to drown them with information at the start.

I have had good results from including designers and the project team in the HUFIT workshops covering activity analyses, context analyses and user analyses. It gives the designers a thorough understanding of the user issues, and includes a short written summary. After this, the designers have been able to carry out their first concept generation phase whilst I have fleshed out a more thorough User Requirements Specification in parallel. In this way, the designer has worked based upon knowledge of users and tasks, but without the constraints of the requirement spec. This allows a creative process for the designer, with (hopefully) the user-task-context information unconsciously steering them in the right direction. When the designers start work on the concept detail, I include the URS as a specification document. I get a feeling that this gives the best of both worlds; it allows the designers to be creative, whilst giving them enough context. If anyone has other variations on this briefing procedure, please get in touch.

Resolution #10: Wear shades

One of the fantastic things about working with designers is not just their creativity but the combination of creativity and strong communication skills. I think that Industrial Designers have an edge over Graphic Designers here, with Architects coming third (architects do however have strong analytical skills and a focus upon structuring information). I always find working with designers a refreshing experience, although the experience varies greatly.

One of the dangers when working with designers is that it is easy to be blinded by the light of design, just because results are presented so well. There is something incredibly seductive about designers' renderings during the first stages of the design process; you become convinced that the solutions are good, just because they are presented so well. Resolution for next year: watch out for things that look sexy, but which will not work. Train yourself to spot the difference between novelty and truly good design.

Conclusion – Less drowning, more waving; The need for Interactive design management

There is a great deal of benefit to be gained from having one person follow user aspects throughout the process, particularly to link across the phases from user focus to technical focus. The increasing number of skills involved in user interface design means that one person cannot and should not have expertise in all of them. Instead, a project needs to hire in expertise within the project, and someone should be responsible for coordinating this. It is therefore natural that a management task is created for interactive systems design.

Design management is a well established discipline within the traditional design world, and the interactive design manager should be modelled upon traditional design management, with added ingredients to cover the human factors and technical implementation aspects. User interface design has become a design discipline and is no longer purely a human factors discipline. We should take the consequences of this and start to learn ow to work together with other disciplines. After all, we all have the same goal.

Who then should the design manager resemble? The skill make-up should be something like is:

- Strong founding in methods for user requirement capture, task analysis, survey design, etc.
- Detailed knowledge of the design process
- Basic knowledge of the human cognitive process and perception
- Marketing and branding expertise

- Industrial and Graphical design knowledge, particularly awareness of different design groups' styles
- Media and Multi-media knowledge
- Knowledge of technical implementation
- Project management skills.

The role of the interactive design manager is to lead the user side of a project, maybe even the whole project, although more likely part of a larger team. The area of responsibility for the design manager is to ensure that adequate user requirements analyses are carried out in a project and that the project has a mandate which is clearly anchored on the user and market side. The design manager should have responsibility for developing a prototype and testing this. Further, he/she should have responsibility for the user requirements specification, the styleguide, the storyboard and to follow these from start to finish, with user testing underway. In addition, he will have responsibility for the user documentation, and for organising necessary training.

As for how many resolutions I will manage to get to grips with, only time will tell. At any rate, I will go out of this year knowing a lot more about what I do not know. Year 2000 problems anyone?

Simon Clatworthy is Research Scientist at Telenor Research & Development, where he has been employed since October 1996. He has worked with command and control systems and systems for public use, and more recently, Internet on the TV. During this time he became interested in the design process, and how best to integrate human factors methods.

e-mail:
simon.clatworthy@fou.telenor.no

Special

First installations of WDM, optical ADM and optical in-line amplifiers on long-haul cable links in Norway

STEINAR M. SVENDSEN

Indisputably, the first optical wavelength division multiplexing (WDM) systems to be established in Norway, appeared almost ten years ago on a handful of regional short-haul routes in Oslo. Those were the glorious days of PDH, one 565 Mbit/s channel was transmitted in each of the two windows, 1310 and 1550 nm. The WDM systems of 1997 imply the introduction of a new era. The test project referred to in this paper has successfully demonstrated transmission of four 2.5 Gbit/s channels over 520 km fibre cable without any need for regeneration, optical equalization or dispersion compensation between the terminals.

Introduction

Early 1997 Telenor AS was offered to perform FVO (Field Verification Office) tests of DSC Communications' first generation of WDM equipment. The original FVO test programme was supplemented, by including additional measurements required for the task to serve as a formal Telenor type approval test. The latter would largely comprise parameters which are listed in the new ITU-T Recommendation G.692, however, without any specific values attached to them yet.

The test programme was carried out partly at the premises of DSC in Denmark, partly as a field trial in Norway. The cable route between Oslo and Bergen, approximately 520 km, was chosen for the field trial. Six optical in-line amplifiers are being used, and no regeneration between the two terminals is required.

The FVO tests commenced on 30 June 1997 in Denmark, and was successfully completed in Norway on 26 August, including a 31 days on site stability test.

The Oslo–Bergen system is now in operation, carrying traffic. Although tested with four channels, provisionally only one is now equipped.

The optical ADM is scheduled for testing first week of December 1997. Following the approval tests, it will be installed on a second WDM route, Gjøvik – Trondheim.

This paper briefly describes the FVO trial routes and the WDM equipment, subsequently presents the main results from the expanded FVO technical report.

Cable route Oslo – Bergen

Table 1 details the amplifier section lengths. On two sections attenuation difference between the two fibres being used exceeds 2 dB, otherwise 0.3 dB or less. Attenuation coefficient per section is 0.23 – 0.29 dB/km, however, one value reaches 0.32 dB/km (Section 1).

It should be noted that the 120 km section from Hønefoss to Gol comprises cable which was installed during the 1985–88 period, when only operation in the 1310 nm window was foreseen. The basic problem of this 'old' fibre type is high mode field diameter and low cut-off wavelength. The remaining part of the route is double-window standard single mode fibre (SSMF - G.652).

Measurements of the 'old' cable type have shown that there will be no noticeable increase in fibre attenuation at -20°C or above. However, below -20°C the attenuation at 1550 nm will increase considerably, depending on individual fibre properties. The performance during the winter months in the Hønefoss–Gol area, amongst the coldest in Norway, will therefore to some extent rely on effective temperature isolation, in particular on snow depth.

Table 1 Section lengths of the cable route Oslo-Bergen

Section no.	Section name	Length (km)	Length (dB)
1	Oslo(Løren)–Gran	75.6	23.9/21.8
2	Gran–Sokna	79.0	21.7/21.7
3	Sokna–Nesbyen	76.1	20.6/20.3
4	Nesbyen–Geilo	71.4	20.6/18.1
5	Geilo–Myrdal	81.7	21.3/21.6
6	Myrdal–Bulken	60.0	13.8/14.0
7	Bulken–Bergen	76.5	18.7/19.0
Total		520.3	

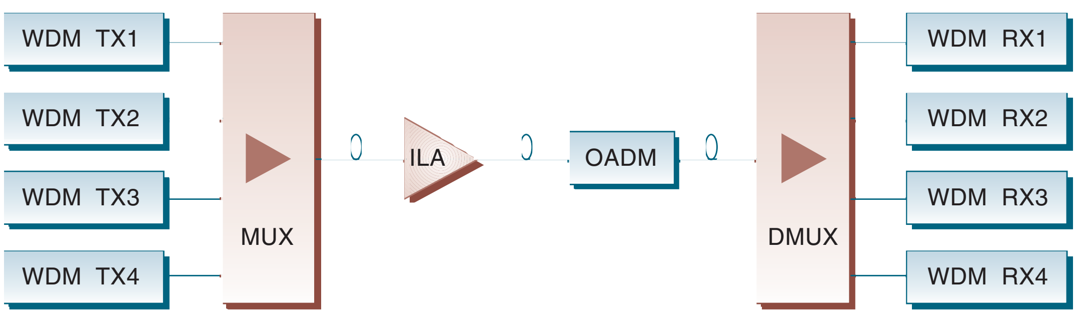


Figure 1 Example of a simple point-to-point WDM system

On the other hand, FVO tests proved that the system margins with respect to attenuation are considerable. In any case, the 'old' cable is considered to be replaced by double-window SSMF.

Table 2 Remote measurements of optical parameters

Sokna				
Direction	**East–West**		**West–East**	
Supervisory system	**NM2100**	**NM300**	**NM2100**	**NM300**
Total input power (dBm)	-12	-12.0	-9.5	-9.5
Total output power (dBm)	+10	+10.0	+9.5	+9.5
Laser bias current (mA)	162	162	142	142
Laser temperature (°C)	25	25	24	24
Pump power (mW)	43	43	40	40

Table 3 Measurements of central wavelengths

Nominal wave-length (nm)	DSC tolerance (nm) Temp./Life	ITU-T range (nm)	Oslo–Bergen (nm)	Bergen–Oslo (nm)
1549.32	± 0.2	1548.67 – 1549.96	1549.7	1549.8
1552.52	± 0.2	1551.88 – 1553.17	1553.0	1553.0
1555.75	± 0.2	1555.10 – 1556.39	1556.2	1556.2
1558.98	± 0.2	1558.33 – 1559.63	1559.5	1559.4

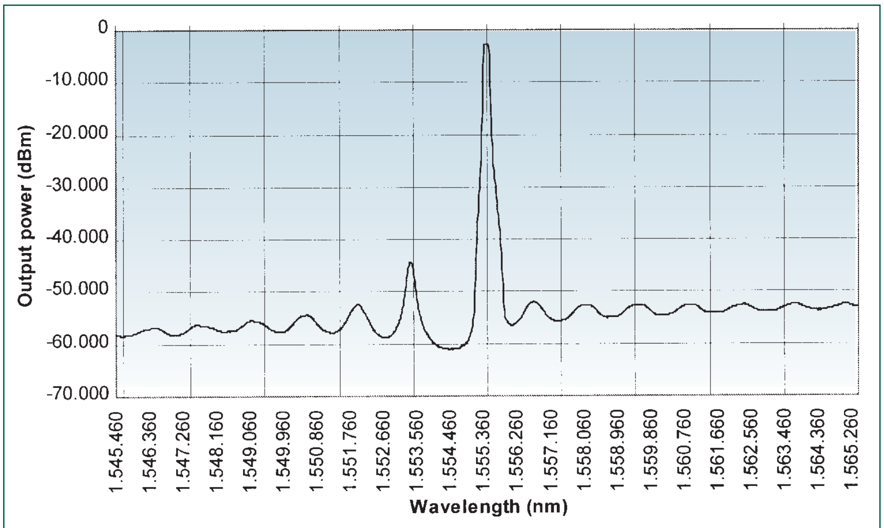


Figure 2 Spectral characteristics of optical transmitter No. 3 in Oslo

The DSC WDM product concept, as delivered to Telenor

The DSC FOCUS WDM is a new product family, to be integrated with the DSC FOCUS AC4 SDH product line. The following system configurations will be supported: WDM point-to-point, WDM add/drop and WDM branching systems.

The first releases feature up to four WDM channels. A network example is illustrated in Figure 1 (only one direction is shown).

- The optical transmitters (TXs) use wavelength selected lasers according to the ITU-T G.692 wavelength grid (400 GHz spacing, 1549 – 1559 nm).
- The optical multiplexer (MUX) couples and amplifies the four channels.
- The in-line amplifier (ILA) operates with a constant total output power. Two alternative optical interfaces are available, 22 dB and 33 dB nominal gain, corresponding to section lengths of approximately 80 km and 120 km, respectively. The 22 dB ILA is a two-way optical repeater, using one separate amplifier for each direction of transmission.
- The optical add/drop multiplexer (OADM) is bidirectional, dropping and adding channels in both directions, using fixed optical filters.
- The optical demultiplexer (DMUX) – combined with the MUX in one single module – just amplifies and splits the incoming WDM signal.
- The optical receivers (RXs) incorporate tuneable optical filters for dynamic optical demultiplexing.

The WDM product family also includes other modules (optical transponders, demultiplexers with fixed filters, optical branching), but these were not tested.

For supervision and management of the WDM elements, the DSC NM2100 network management system can be used, as well as NM300-WDM as a local craft terminal.

Trial routes configurations

For the trial routes at DSC and on site in Norway, four 2.5 Gbit/s channels (double sets of optical transmitters/ receivers) were equipped. At both locations also six 22 dB ILAs were being used.

The link at DSC was equivalent to the Oslo–Bergen route, except that in one direction of transmission optical attenuators were inserted instead of fibre drums between amplifiers.

The DSC system is capable of coping with a fibre dispersion of up to 10,000 ps/nm, corresponding to a total system length of approximately 540 km.

Test of remote supervisory system

A set of five optical parameters per direction, from each ILA and each MUX/DMUX, can be measured.

An example is shown in Table 2. The ILA of 'Sokna' was at first checked at DSC, using both NM2100 (the network management system, connected to MUX/DMUX in 'Oslo' via the Terminal Server interface box) and NM300 (local craft terminal, connected to the ILA of 'Gran'). Obviously, the results are corresponding.

NM2100 is also capable of forcing-off/switching-on individual lasers, LOS will be reported to both NM2100 and NM300.

Central wavelengths

Test instrument used: W&G OMS-100 Optical Measuring System (calibrated).

All measurements were performed in Oslo. Results are given in Table 3.

All values measured are relatively high, above nominal figures. They are outside DSC tolerance figures, however, within the ITU-T recommended range. The corresponding measurements previously carried out at DSC, using the optical spectrum analyzer, were close to nominal wavelengths. Obviously, test results may differ more than expected when using different test instruments, although both calibrated.

Spectral characteristics of optical transmitters

Test instrument used: Ando AQ-6315A Optical Spectrum Analyzer.

As laser modulation bandwidth is in the order of 0.02 nm, the -20 dB width could not be measured using the optical spectrum analyzer.

For the scanning a resolution of 0.2 nm and Mode 'High 1' was used. (This mode suppresses noise and performs a more correct measurement of the optical signal. Measurement duration is also convenient.)

All optical transmitters of the 'Oslo' terminal station were measured at DSC. The spectral characteristics of transmitter No. 3 is shown in Figure 2, as an example.

Test results 'Oslo TX3':

Central wavelength 1555.440 nm
Power level -2.67 dBm

Highest side mode 1553.380 nm
Power level -44.27 dBm

Wavelength difference -2.060 nm
Power difference -41.60 dB

Vacuum measurements were also performed, resulting in a central wavelength of 1555.88 nm, i.e. 0.44 nm higher. (The ITU-T Recommendation G.692 refers to vacuum, nominal value and range are shown in Table 3.)

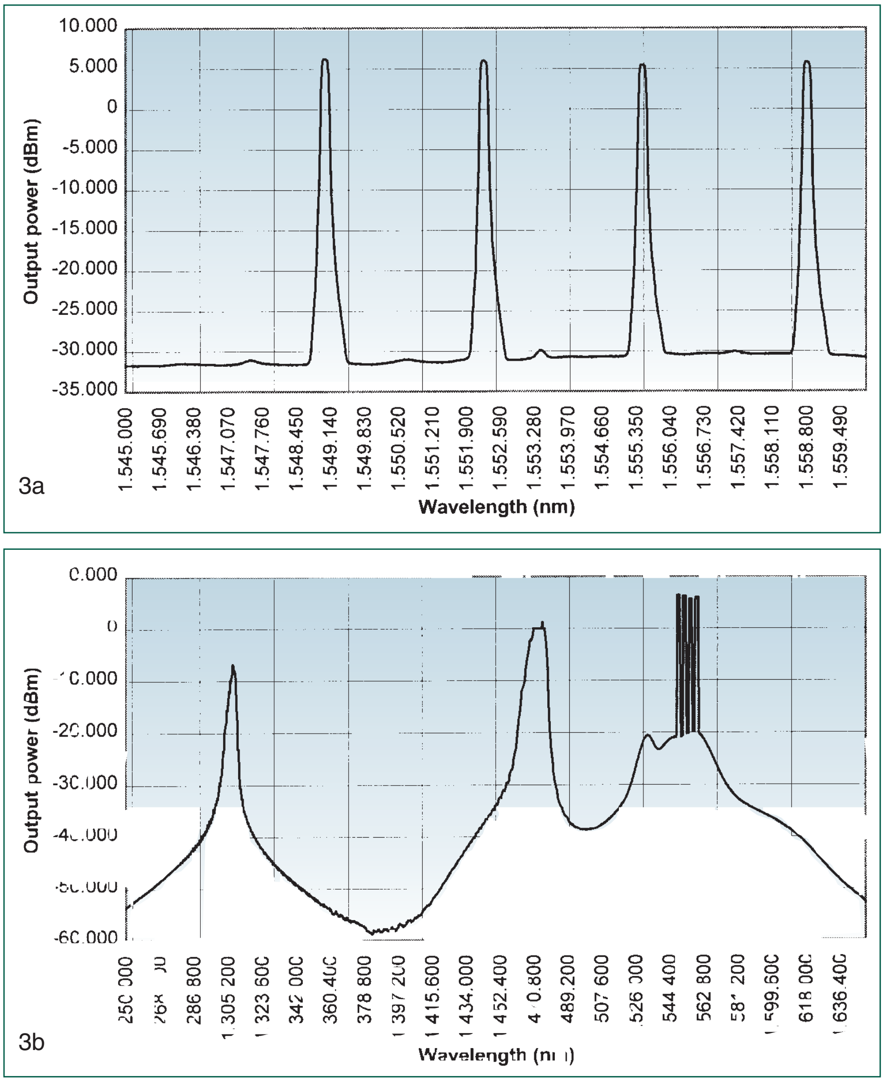


Figure 3a/3b Spectral characteristics at MUX output in Oslo

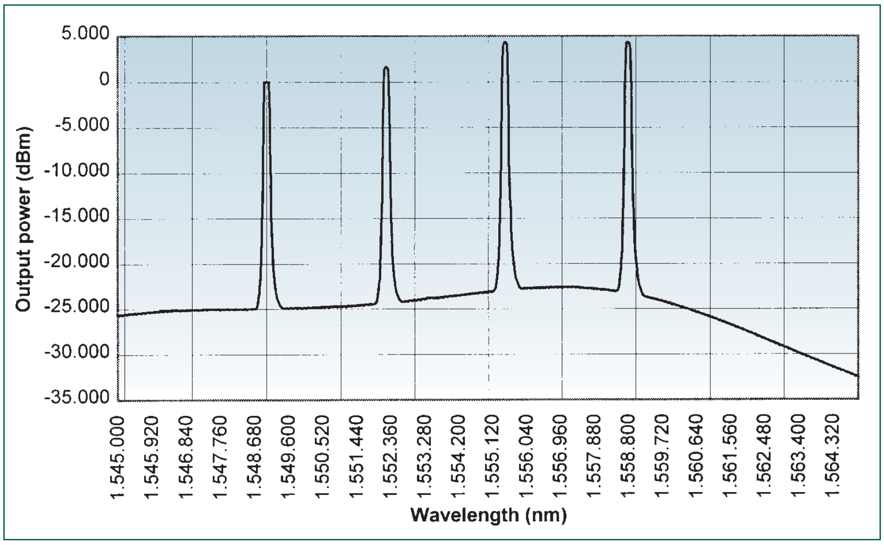


Figure 4 Optical spectrum at Bulken ILA output

Table 4 Individual channel power level (dBm) at ILA outputs and Bergen DMUX output

ILA name ILA no.	Gran 1	Sokna 2	Nesbyen 3	Geilo 4	Myrdal 5	Bulken 6	Bergen DMUX
1548.98 nm	2.69	2.22	1.34	1.52	0.73	0.82	-12.04
1552.24 nm	2.78	2.60	1.90	2.07	1.14	1.59	-10.81
1555.44 nm	3.29	3.55	3.56	4.06	3.30	4.29	-7.81
1558.76 nm	3.64	3.85	4.00	4.25	3.18	4.29	-7.82
Max ch. power diff.	0.95 dB	1.63 dB	2.66 dB	2.73 dB	2.57 dB	3.47 dB	4.23 dB

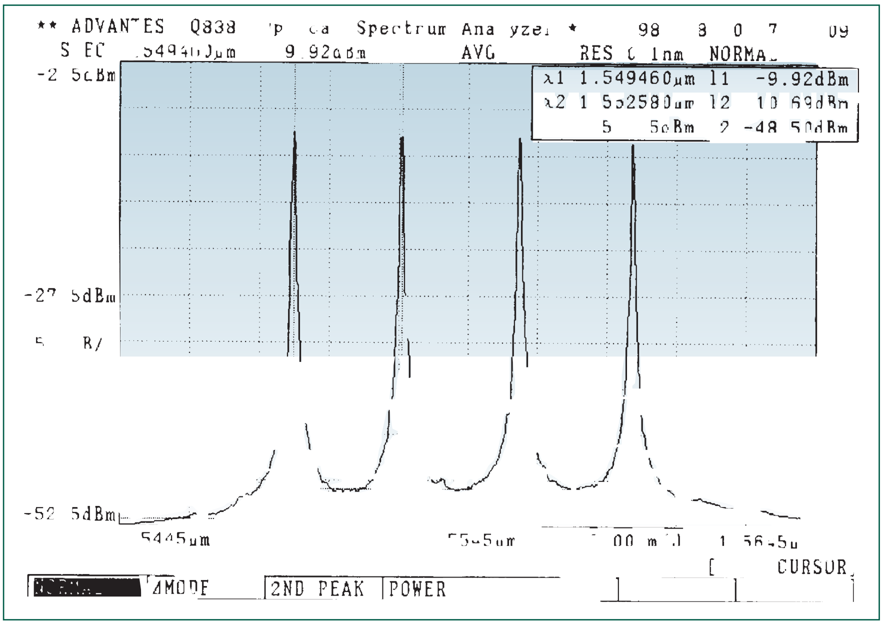


Figure 5 Optical spectrum at Oslo MUX/DMUX Test Out (direction Oslo–Bergen)

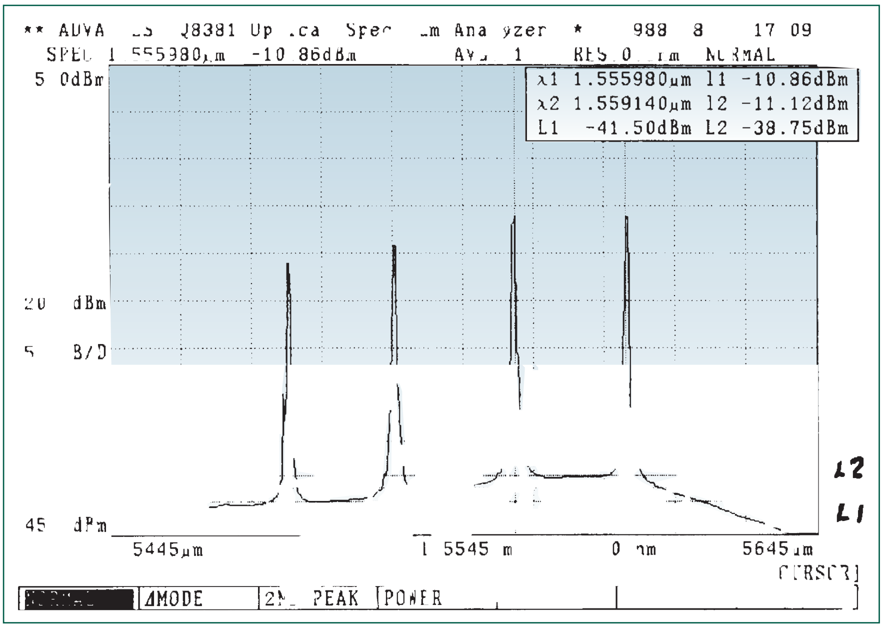


Figure 6 Optical spectrum at Oslo MUX/DMUX Test In (direction Bergen–Oslo)

Side mode suppression ratio for the four transmitters in Oslo were found to be:

Oslo TX1: 45.72 dB
Oslo TX2: 47.94 dB
Oslo TX3: 41.60 dB
Oslo TX4: 46.10 dB

DSC requirement: SMSR = 30 dB
ITU-T G.957 recommendation for STM-16: SMSR = 30 dB.

Central wavelengths are all lower than the values later to be measured in Oslo (using a different test instrument).

Optical spectrum along the route

The optical spectrum was recorded at MUX output, all ILA outputs, DMUX output and DMUX test output. For practical reasons this test was carried out at DSC, and in Norway only repeated in Oslo MUX/DMUX.

Oslo MUX output, as measured at DSC, is shown in Figure 3a (1545 – 1560 nm) and 3b (1250 – 1650 nm). In Figure 3b one will see, from left to right: the OSC (optical supervisory channel) at 1310 nm, the optical pump power at 1480 nm and finally, the four traffic channels.

The per channel power at Oslo MUX output was measured to be (using the optical spectrum analyzer):

1548.980 nm: +6.78 dBm
i.e. 4.76 mW
1552.240 nm: +6.32 dBm
i.e. 4.29 mW
1555.440 nm: +6.12 dBm
i.e. 4.09 mW
1558.760 nm: +6.38 dBm
i.e. 4.35 mW
In total +12.40 dBm (17.37 mW)
i.e. in total 17.49 mW

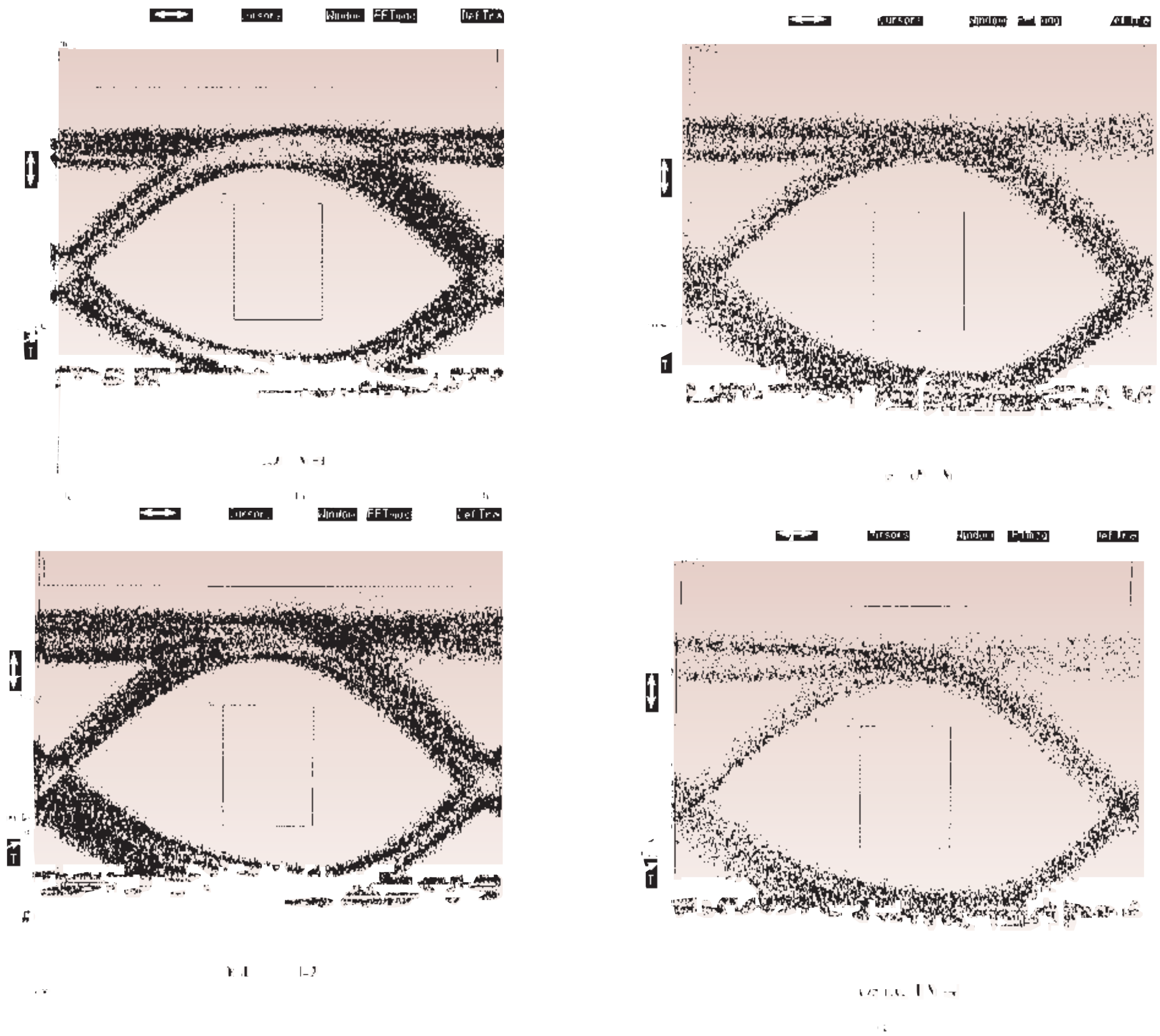


Figure 7 Eye diagrams at optical transmitter outputs in Oslo

Maximum channel power difference: 0.66 dB. SNR > 35 dB.

Moving from Oslo to the sixth ILA, Bulken, the optical spectrum has changed some, as one might expect. Please refer to Figure 4.

Channel power levels, at the output of each ILA, are shown in Table 4.

The spectrum diagram at Bergen DMUX output indicates an SNR > 24 dB. This means that from MUX output in Oslo to DMUX output in Bergen SNR has decreased 10 – 13 dB.

As can be seen from Table 4, maximum channel power difference increases along the route (albeit irregularly), reaching a value of 4.23 dB in Bergen. This power difference is no problem to handle (10 dB can be accepted).

The optical spectrum measurements performed on site in Oslo, made use of an Advantest Q8381 Optical Spectrum Analyzer.

Results are shown in Figure 5 (Oslo MUX Test Out – i.e. direction Oslo–Bergen) and Figure 6 (Oslo DMUX Test In – i.e. direction Bergen–Oslo). From these results one may conclude:

SNR is clearly better than 35 dB for all outgoing channels, in the area 25.5 – 28 dB for all incoming channels.

Maximum channel power difference is 5.2 dB, i.e. slightly more than what was measured at DSC, direction Oslo-Bergen.

Eye diagram and extinction ratio

Eye diagrams of all optical transmitter outputs (measured in Oslo) are shown in Figure 7.

All values of extinction ratio (ER) range from 9.1 to 14.0 dB. According to DSC product specification, the minimum value for ER should be 8.2 dB. (This figure is in agreement with the ITU-T requirement for 1550 nm V-type STM-16 interfaces for

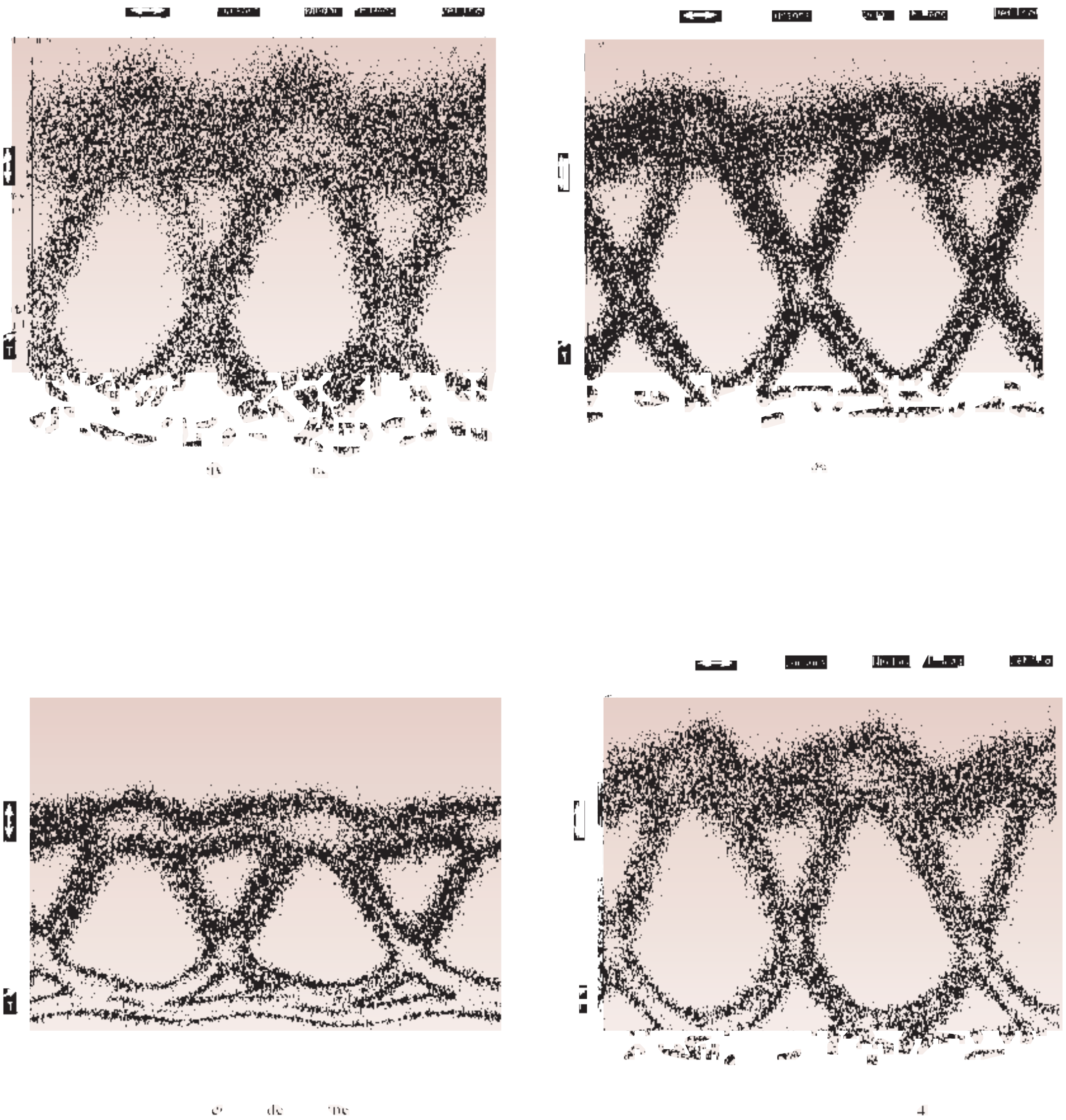


Figure 8 Eye diagrams at optical receiver inputs in Oslo

single channel systems. There is not yet any requirement in G.692.)

Eye diagrams of all optical receiver inputs (measured in Oslo) are shown in Figure 8.

The four eye diagrams are noticeably different, as also was indicated on transmit side. Obviously, the Oslo–Bergen route seems to be approaching the maximum systems length.

System margins

Attenuation margins were tested at DSC, using optical attenuators instead of cable on all the seven amplifier sections. Various combinations of reduced section attenuation were tried, e.g. attenuation in three sections was set to 10 dB simultaneously. Test time varied from 3 minutes to 15 minutes, no bit errors were reported.

A BER of 10^{-12} corresponds to approximately 0.56 errors per hour, or 13.44 errors per 24 hours referring to STM-1. The measurements above are therefore indicative only.

Amplifier section attenuation was also increased. Section 7 could accept 17 dB additional attenuation, no errors in a 7 minutes test period. Section 1 could accept 16 dB additional attenuation, no errors in a 3 minutes test period. Section 5 could accept 9 dB additional attenuation, no errors in a 3 minutes test period.

Finally, attenuation was increased simultaneously in 4 sections, numbers 1, 2, 5 and 7. The total increase corresponded to 35 dB beyond the nominal maximum (22 dB x 7 = 154 dB). No errors were observed during the test period. The margins seem to be pretty safe with respect to attenuation.

To investigate dispersion penalty, a simplified test was carried out in Norway. It was concluded that penalty caused by cable dispersion was negligible.

31 days stability test

The Oslo–Bergen system was looped in Bergen on STM-1 level and supervised for one month of operation. No errors occurred during this period. This result may seem satisfactory.

Other tests

Automatic laser shutdown and restart functions (ALS and ALR) were verified. Output power from optical transmitter alternated between 'on' (-2.5 dBm) for 2 seconds and 'off' for 70 seconds after disconnection. When applying manual restart, 2 seconds pulses and 90 seconds pulses (for measurement purposes) respectively, were checked.

Basic alarm test was performed.

Some module tests were left out, as a test report prepared by DSC contained the necessary information, and there did not seem to be a reasonable need for repeating them.

Conclusion

An overall evaluation of the test results easily concludes that DSC FOCUS AC4 Release 2.5 WDM equipment can be approved for application in Telenor's network. Performance has been proven to be satisfactory, partly beyond expectations, and key system margins seem to be more than adequate.

Acknowledgements

The author wishes to acknowledge all who participated in the test programme,

- from DSC Communications A/S: Aiguo Li, Carsten G. Jørgensen, Martin Seerup Nielsen, Ole M. Olesen
- colleagues from Telenor AS: Steinar Bjørnstad, Ottar Flydal, Irene Kaasbøll and Jon Tjøgersen.

Also appreciated is the kind permission granted by DSC Communications to publish product and measurement material contained in this paper.

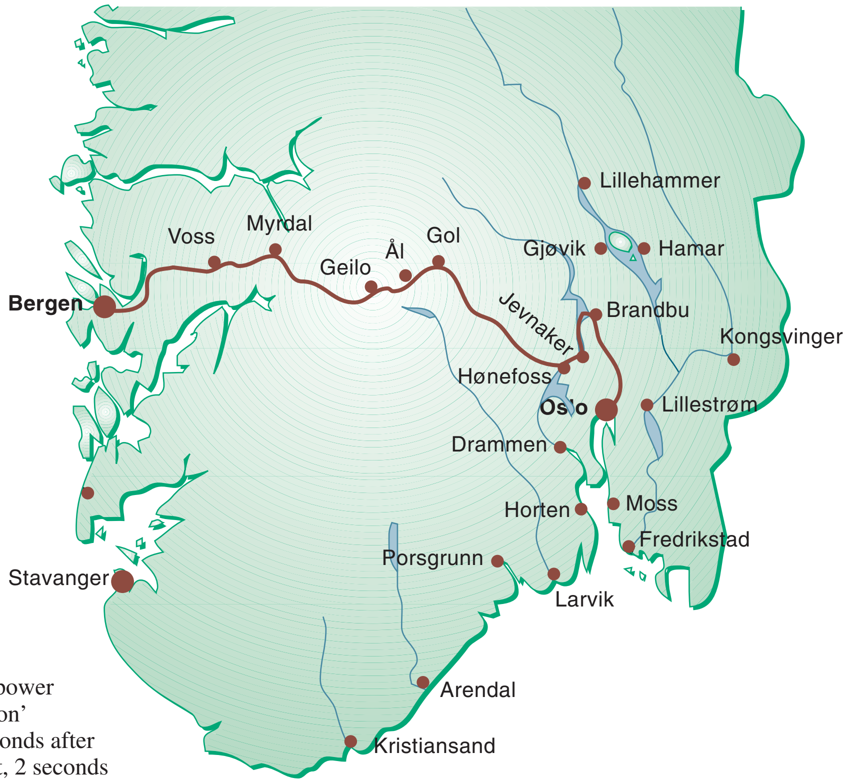


Figure 9 Cable route from Oslo to Bergen

Steinar M. Svendsen is Chief Engineer at Telenor Nett AS, Infrastructure Division. Since 1983 his major area of interest has been optical line transmission systems for the transport network. He is currently working within the SDH project of Telenor.

e-mail: steinar-m.svendsen@s.nett.telenor.no

Status

International research and standardization activities in telecommunication

Editor: Per Hjalmar Lehne

Introduction

PER HJALMAR LEHNE

In this issue of the Status section we take a closer look into some projects from the ACTS (Advanced Communications, Technologies and Services) programme. ACTS has a total budget of 670 MECU, or 5 % of the total budget for the European Union Fourth Framework Programme. The programme is divided into six technical domains and one supporting, or horizontal, domain:

- Interactive Digital Multimedia Services
- Photonic Technologies
- High-Speed Networking
- Mobility and Personal Communication Networks
- Intelligence in Networks and Services
- Quality, Security and Safety of Communication Systems and Services
- Horizontal Actions.

The ACTS programme is probably the most important driving force in Europe into the future of telecommunications technologies.

Telenor is currently participating in 17 projects in ACTS, additionally there are Norwegian participation in three more projects. The projects are listed in the table. Four of the projects are presented here.

First, Mr. Agne Nordbotten describes the *CRABS* project (Cellular Radio Access for Broadband Services). This project focuses on interactive services operated at 40 GHz. Trials will be performed in five European countries. It is expected to form a foundation for future system recommendations. In this project, Telenor has the leadership as *Prime Contractor,* with Mr. Nordbotten as project leader. In addition to describing the project objectives, he also shares some of his experiences of having the leadership and the benefits for Telenor.

Mr. Torodd Olsen has, together with the other partners, contributed with a description of the *OPEN* project (Optical Pan-European Network). This project focuses on the possibilities of using an overlay network based on optical fibres and transparent optical cross-connects (OXC). Field trials will be performed between Norway and Denmark and between France and Belgium. The main objective is to assess the feasibility of the concept. Telenor's role in the project is also presented.

Multimedia, i.e. joint use of text, audio and video to provide information, or as a communications means, is a growing field. This is made possible by the merging of advanced computing and information processing and of broadband telecommunications technologies. In the near future, multimedia services will also be available when mobile. Mr. Robert Danielsen presents the *MoMuSys* project (Mobile Multimedia Systems) which aims at the development of mobile multimedia technology. The project is very strongly linked up to standardisation work in ITU[1] and ISO[2]. He describes the results up to now and also has some thoughts about how the work can be continued.

Finally, Ms. Vendela Paxal presents the *SINUS* project (Satellite Integration into Networks for UMTS[3] Services). This project's main objective is to define techniques for the integration of the satellite component into UMTS. Special focus is on the interworking with the terrestrial component. As in all ACTS projects, trials play an important role also here.

Table 1 ACTS projects with Norwegian participation

Interactive Digital Multimedia Services	
CRABS	Cellular Radio Access for Broadband Services
MAESTRO	Maintenance System based on Telepresence for Remote Operators
Mobility and Personal Communication Networks	
MoMuSys	Mobile Multimedia Systems
SAMBA	System for Advanced Mobile Broadband Applications
SINUS	Satellite Integration into Networks for UMTS
STORMS	Software Tools for the Optimization of Resources in Mobile Systems
High-Speed Networking	
ASICCOM	ATM Switch for Integrated Communication Computation and Monitoring
CA$HMAN	Charging and Accounting Schemes in Multi Service ATM Networks
EXPERT	Platform for Engineering Research and Trials
JAMES	Provision of European Networking Facilities (PEN)
NICE	National Hosts Interconnection Experiment
REFORM	Resource Failure and Restoration Management in ATM-based IBCN
SMASH	Storage for Multimedia Applications Systems in the Home
TRUMPET	TMN's Regulations and Multiple Providers Environment
Photonic Technologies	
BLISS	Broadband Lightwave Sources and Systems
MEPHISTO	Management of Photonic Systems and Networks
OPEN	Optical Pan-European Network
Quality, Security and Safety of Communication Systems and Services	
RETINA	An Industrial-Quality TINA-Compliant Realtime DPE
Horizontal Actions	
INFOWIN	Multimedia Information for National Hosts
OPTIMUM	Optimized Network Architectures for Multimedia Services

Per Hjalmar Lehne is Research Scientist at Telenor Research & Development, Kjeller. He is working in the field of personal communications, with a special interest in antennas and radio wave propagation for land mobile communications.

email address: per.lehne@fou.telenor.no

1 *International Telecommunications Union, a UN body.*
2 *International Standardisation Organization.*
3 *Universal Mobile Telecommunications System.*

The ACTS Project AC215 CRABS; Cellular Radio Access for Broadband Services

AGNE NORDBOTTEN

This article gives a presentation of the ACTS project AC215 CRABS which focuses on the development and demonstration of a broadband radio system for interactive services operated at 40 GHz. The project covers service and user aspects, architecture and systems studies including extensive simulations as well as studies of propagation issues to determine propagation type QoS criteria and coverage models and equipment studies including technical trials. The project is highly trials oriented and trials have now started in five European countries. The final results of the project are expected to form the foundation for recommendations of a future system.

1 Introduction

The main objective of this project is to develop and demonstrate a broadband radio access system with interactivity operated in the frequency range 40.5 – 42.5 GHz which is the frequency range allocated in most European countries. The results from these trials in combination with systems and propagation studies as well as evaluation of technology, equipment and user and service requirements, will form the basis for the final systems recommendations. The project is thus focusing on extensive technical, user and service trials performed at different locations in five European countries, systems studies and simulations, establishment of propagation and coverage models to be used for network planning as well as evaluation of results and recommendations.

The project started July 1996 and has a duration of 30 months with a total effort of altogether 1017 man months divided among 9 partners and 3 associate partners as listed in Table 1.

This paper presents the philosophy and objectives of the project, the main tasks addressed in the different workgroups, the objectives of the trials and some preliminary conclusions from the first trials.

Table 1 Partners of the CRABS project

Code	Partner name	Partner roles
P01	Telenor R&D, Norway	Prime contractor, WG leadership, Trial site
P02	Philips Broadband Networks, UK	Ind. partner, WG leader
AP02A	Philips Research Laboratories, Redhill, UK	Associate partner P02, Res.org.
AP02B	Philips SpA, Monza, It	Associate partner P02, Res.org.
P03	Rutherford Appleton Laboratories, UK	WG leadership, Propagation trials
P04	Joanneum Research, Austria	Research organisation associated with the University of Graz
P05	TESTCOM, Czech Rep.	Research Organisation, Trial site
P06	RAI, It	Broadcaster, Trial site
AP06A	Politico di Milano, It	University, associated with P06
P07	EUROBELL, UK	Cable operator, Trial site
P08	DEMOKRITOS, Greece	Research organisation, Trial site
P09	THOMSON CSF, Fr	Industrial organisation

2 Background and basic systems aspects

Originally, the frequency range 40.5 – 42.5 GHz was planned to be used for broadcasting or redistribution of broadcasting, a wireless cable system. As the cost of cable provision was estimated to be between 20 – 40 ECU per metre, depending on local conditions, a significant number of rural and suburban homes with long feeder connections could be denied broadband access to interactive services. However, the event of digital broadcasting by satellite opened for new broadcasting related services including interactivity and integration of data and telecommunication services. The growth of Internet traffic clearly demonstrated the possibilities of new interactive services. It became widely recognised that broadband interactive connections to the homes and to smaller enterprises and organisations would be required and that a radio based solution could be very attractive and flexible when sufficient bandwidth was available.

The project discussions started at the very beginning of digital broadcasting by satellite and during the first Internet wave. It seemed obvious that services based on a combination of broadcasting, interactivity, data exchange and new

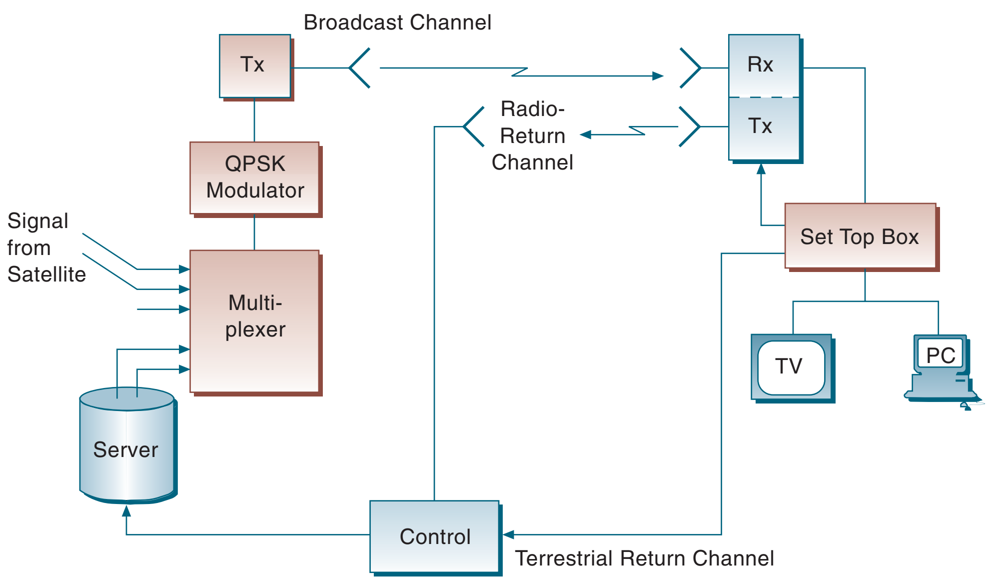


Figure 1 Sketch of basic system

very flexible telecom services, could have an almost enormous growth potential into the beginning of the next century. It was felt that a broadband system based on the new digital broadcast technology, including data, for the downlink and a flexible use of communication technology for the uplink, would offer the possibility of creating a cellular access system for real broadband services in the local domain. A systems solution along these lines is outlined in Figure 1.

The solution is based on the application of the MPEG2/DVB technology used for satellite broadcasting including data of any kind. Thus, also QPSK modulation is being used in order to be able to use the digital Set Top Boxes developed for satellite reception. At the receiver site the 40 GHz signal is down-converted from 40 GHz to the intermediate frequency range 950 – 2050 MHz of the satellite receiver. Reception is presently based on this type of receivers in combination with PCs. It is very important that the user interfaces required for services using different networks are identical to the extent possible. At the transmitter side information signals are multiplexed on to MPEG2 transport streams and transmitted in the frequency range 40.5 – 42.5 GHz. The signals may contain retransmitted TV signals, data obtained from servers through interactivity like Internet services, and on-line services like conferencing. Interactivity may be obtained through available terrestrial technology like ISDN, DECT, or radio relay systems. At the end of the project an in-band return channel at 40 GHz will be included.

3 The partners and their roles

The basic project ideas originated in a small group of participants with three of the final partners as dominant contributors. When the structure was more or less worked out the need for additional partners was discussed from both a technical and political point of view. The decision was to make a winning proposal and it was known that the competition within this field would be very strong and that several proposals were expected. The proposal became a winner with the partners listed in Table 1.

This clearly illustrates the broad and strong combination of operators in the project with Telenor representing the telecom, cable and satellite sides, EUROBELL is a cable operator interested in investigating the possibilities of employing broadband radio technology to extend their existing services and to add new flexible interactive services. The broadcaster RAI is looking for new activities in the local domain and is in this project testing educational aspects, while TESTCOM has the much broader view of introducing digital services as soon as possible in the Czech republic. For DEMOKRITOS the introduction of interactive services is the main objective.

Industry is represented by Philips and THOMSON CSF, Rutherford Appleton Laboratories is recognised as the leading institute on propagation issues in Europe. Joanneum Research is known for contributions to satellite communication through ESA, leadership of COST projects and recognised competence in the field of transmission protocols. Politecnico di Torino, associated with RAI, are specialists on microwave and optical technology.

Thus, the project is well composed with regard to the work to be done.

4 The CRABS trials

Since ACTS is a program focusing on trials and demonstrations, this is also a central part of CRABS with trials being performed in five different countries on the premises of the different partners. While most ACTS projects have more or less centralised trials which very often are located at National Hosts, the CRABS trials are distributed throughout Europe and there is a local motivation behind each trial site depending on geographical location, cultural, economical and technical background, type of operator involvement and level of expectation. Within the project it is felt that this is the best way to create an interest for and to test new services.

Regarding the technological background represented by factors like the density of PCs in the homes and the volume and growth of interactive services, there is a strong gradient from the north to the south of Europe, with the Nordic countries in the lead.

Different types of services have been chosen for the trials, rebroadcasting of digital satellite programs, video on demand, Internet, information services, conferencing and educational services to mention some.

Interactivity is definitely an important issue and the project has, at least from the start, adopted the interactive reference model developed by the ACTS projects INTERACT and DIGISAT. This model is shown schematically in Figure 2.

The downlink channel is a broadcast channel or point to multipoint channel containing the information addressed to groups and individuals. The coverage range is typically 3 – 5 km with an opening angle of 50 – 60°. The uplink, interactive or return channel is a typical communication channel where different technologies may be used depending on availability and the needs of the user.

The CRABS trials are divided into two phases running for approximately one year each. In phase 1 available terrestrial technology like ISDN, DECT and radio relay technologies are being used depending on trial site, while phase 2 will also include in-

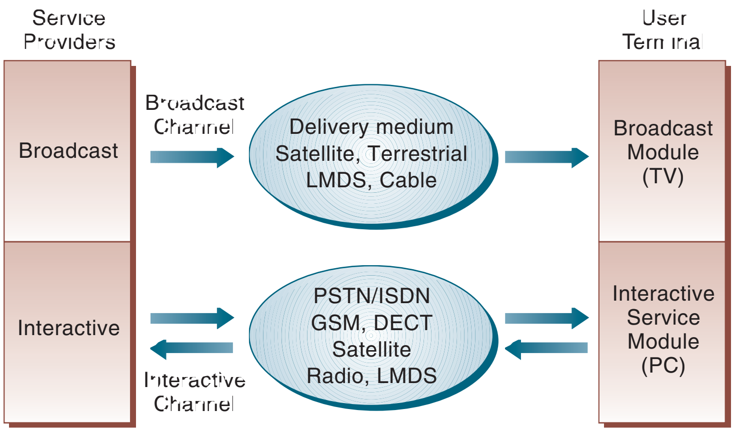


Figure 2 Interactive reference model

band radio return at 40 GHz. Telenor R&D is using ISDN connections for the first phase concentrating on testing the broadcast channel and interactivity in itself with a high capacity delivery channel.

At the trial site in Kent, UK, the cable operator performing the trials is hoping to extend the business potential and a radio path offering high capacity is wanted in order to offer broadband services to smaller enterprises and schools. The Telenor trial at Kjeller is directed towards five selected users/user groups. The interactive channel during phase 1 is ISDN based, and an in-band RF channel will be tested in phase 2 starting next year. In Italy RAI is performing trials/demonstrations at different locations in Venice, Turin and Rome. The Rome trials are mainly educational. The Greek trials are directed towards tourist information purposes. Through connections with databases local information is obtained or connections can be established into Internet where local newspapers etc. from home countries may be obtained.

The trials are expected to give:

- Information on equipment needs
- Variations of QoS with meteorological/topological conditions
- A possibility for testing of both architectural models and coverage models
- A background for contributions to standardisation of systems and services
- A testbed for ideas on how broadband services will develop in Europe
- The views from different operator groups and different locations around Europe regarding new broadband services.

So far the results look promising. Trials are now started in all five trial countries. The main obstacle up to now has been lack of stability in parts of the transmitter equipment. This has to be improved. The studies for further development of the system as well as the propagation studies are giving good results.

5 The role of prime contractor

In this project Telenor has also had the role of prime contractor. It was not the intention during the planning period, but suddenly the situation changed, and the challenge was accepted.

This role definitely takes a lot of effort, and it is very different from a local project where all the participants are well known and the project leader can organise the project with that as a background. In an international project it takes some time to obtain such background knowledge, and at that time the different activities in the project are being addressed, and not necessarily in an optimised way. The creation of motivation in such a project can definitely be a difficult task. In the CRABS project we have been lucky as we have so far avoided destructive and lengthy discussions on minor topics.

The prime contractor also has the responsibility of representing the project towards Brussels, which is an advantage at least when the results from the project are reasonable good. It gives the possibility for active participation in domain meetings and workshops where the flow of information from other projects is considerable, a forum for information exchange and concentrated watcher activities. Presentations of the project also represent possibilities for international contact and recognition.

6 Conclusion

The participation in the CRABS project has given us valuable experience in leadership of relatively large international projects. Access to the international research community in this way is very challenging from a professional point of view. The interest for work on broadband radio access has been growing during the months the project has been running, a fact which is also reflected in the increasing interest for the project. It is expected that the last 15 months of the project will show good and interesting results from trials, systems and propagation studies and that these results will contribute heavily to the development of radio based broadband interactive services. The time from trials to real services seems to be short in this field. Except for some equipment problems at the start, the project got up to a proper speed in a short time and is now moving according to the intentions.

Agne Nordbotten is Research Manager Satellite and Radio Systems, Telenor Research & Development, Kjeller, and Project Coordinator for AC215 CRABS.

email address:
agne.nordbotten@fou.telenor.no

AC066 OPEN – Optical Pan-European Network

TORODD OLSEN

1 Main objectives

The ACTS project OPEN aims at the study of the feasibility of an Optical Pan-European overlay Network (OPEN), interconnecting major European cities by means of a mesh of high-capacity optical fibre links cross-connected through transparent photonic nodes. The network is built on an existing fibre infrastructure, and offers an upgradable capacity. As can be seen in Figure 1, the scope of OPEN addresses the transport in the optical transmission medium layer and the optical path sub-layer. The potential capacity of each fibre link will be upgradable at least up to 40 Gbit/s, each channel supporting STM-16 SDH or higher data-rate transport service. The proposed approach relies on the extensive use of wavelength division multiplexing (WDM) for both transmission and routing purposes, but the interfacing with high-speed trunks carrying optical time-division multiplexed (OTDM) channels are considered as well. The goal of the project is to assess by modelling, laboratory experiments, and field implementation of a photonic network prototype, the feasibility of the OPEN concept [3,4].

2 Network concept and architectures

The advantages brought by the use of WDM for transmission and routing within a network like OPEN are numerous. First, this technique allows for an increase of capacity of each link, without resorting to the use of ultra fast optical technologies. In addition, this increase of capacity is accompanied by a higher resistance to both group velocity and polarisation mode dispersion of the line fibre, which limit the transmission distance through the individual channel rate. Therefore, WDM provides non-regenerated span x throughput products that are substantially higher than other techniques. Second, WDM allows for an enhancement of flexibility of routing techniques, by the use of the wavelength dimension. In this context, all-optical WDM routing techniques in the nodes offer the possibility of implementing simple network configuration, protection, restoration and management schemes. Finally, and perhaps most importantly, all-optical WDM routing techniques provide such a network with a high degree of transparency: any type of service, with any bitrate, and adopting any digital frame format, can be handled and transported from one end of the network to the other with the same optical equipment, avoiding the need of electrical regeneration in the nodes.

In terms of network analysis, the project has established the full description of the network, including the overall dimensioning, the evolutionary path from the existing infrastructures, the operation and management aspects and the protection schemes, so as to propose a sound vision of the various steps leading to the implementation of the OPEN network in a medium-term and in a longer-term perspective. Four topologies have been studied for the network, as depicted in Figure 2. The first one is the flat topology, which optimises the position of the OXCs with respect to traffic density and capacity. The second one is the partitioned topology, which seeks to partition the network into WDM islands, which may then be interconnected with OTDM high-speed trunks. The hierarchical topology subdivides the network into layers in a way that the number of hops required in the interconnection of the two nodes is minimised. In addition, a fourth approach, termed a modular/hybrid topology, is also being studied. Furthermore, the project is developing an end-to-end transmission model based on node and link modelling activities.

3 Enabling technologies

In the OPEN network concept, all optical WDM routing is implemented by means of multiwavelength optical cross-connects (MOXCs), located at major European cities and interconnected via WDM links. All-optical WDM technologies are used in this concept both for transmission between the network nodes and routing through MOXCs, exploiting exclusively optical regeneration and reshaping of incoming data. The WDM components for transmission are flat-gain optical amplifiers and wide-band dispersion compensating devices. The main requirement on these components is a wavelength-insensitive operation in the entire spectral band of the multiplexed channels. On the other hand, the principal functional blocks of the MOXC are a multiwavelength space switch, a tuneable optical filter, an all-optical wavelength converter and a flat-gain optical amplifier. Figure 3 shows the architecture of the OPEN MOXC, which is based on the broadcast-and-select principle. The incoming multiplexed signal from each port is split as many times as the number of output wavelength channels. For each output channel, the correct input port containing the relevant input channel is selected at the switching stage. At the tuneable filter stage this channel is selected from the input multiplexed signal and, if necessary, its wavelength is translated before multiplexing into the outgoing fibre.

The all-optical wavelength converter is a non-linear device that provides some signal reshaping such as extinction ratio enhancement, limitation of noise accumulation and spectral

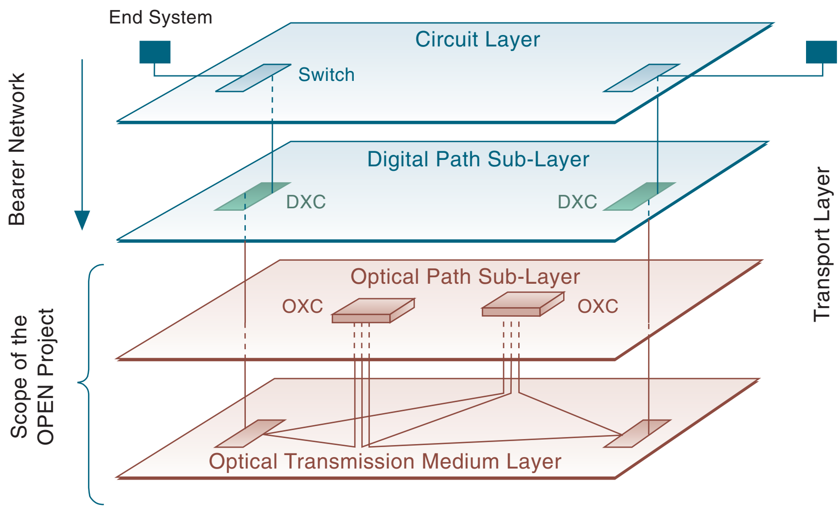


Figure 1 OPEN network concept

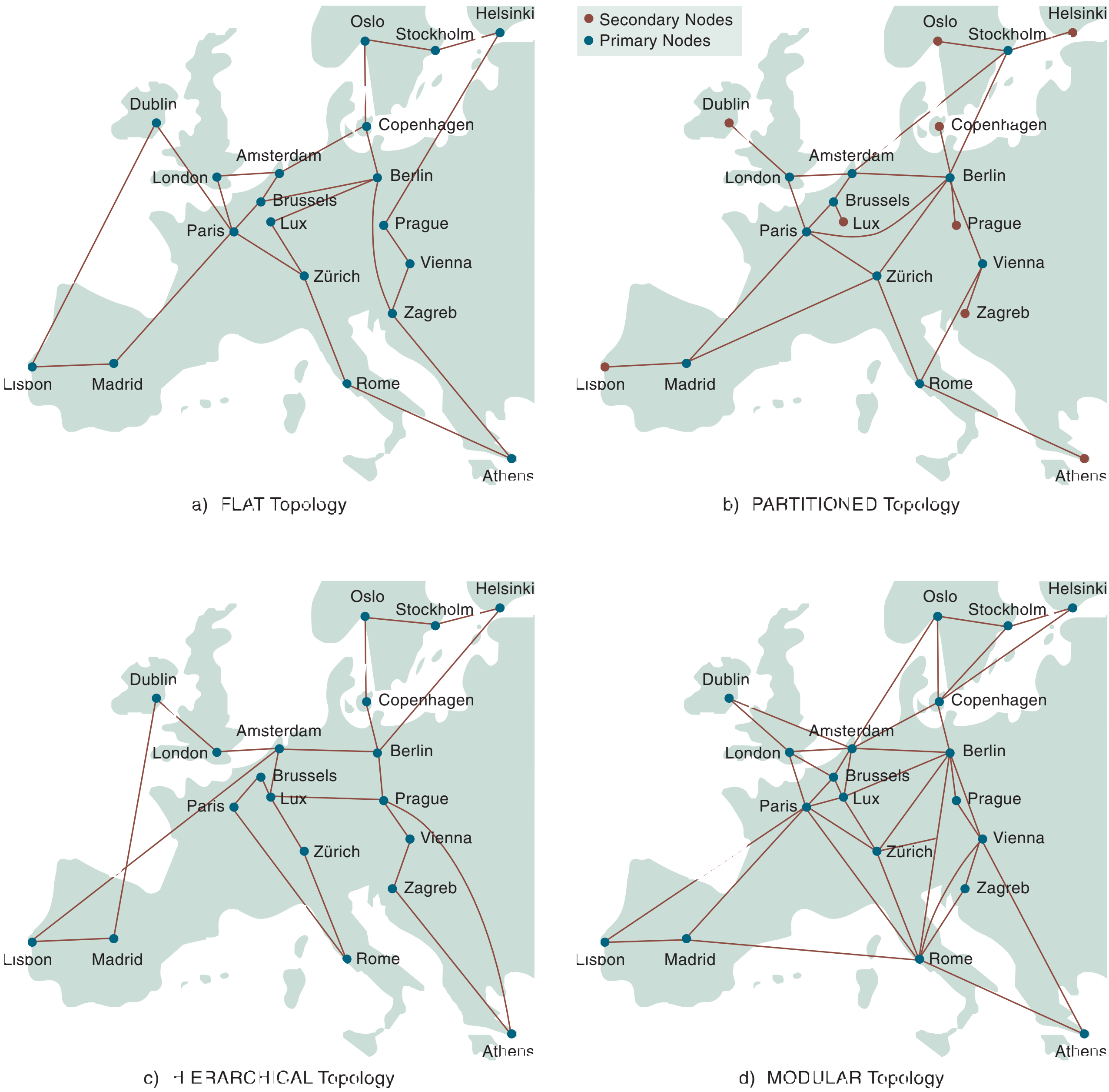


Figure 2 The four topologies being studied in OPEN

cleaning, allowing to overcome several physical limitations (like noise accumulation, cross-talk in switches and filters, filter bandwidth narrowing, chromatic dispersion and non-linear effects in the transmission spans) that would otherwise drastically limit the achievable network expansion. The OPEN cross-connect is implemented using some advanced optical components, among which are a semiconductor-based optical switch optimised for multiwavelength operation and an all-optical interferometric wavelength converter.

The laboratory experiments carried out by OPEN are using recently developed advanced optical components. The project has demonstrated successfully the transmission of 16 x 10 Gbit/s signals over 531 km of standard fibre, using dispersion management [1]. These experiments featured Dispersion-Compensating Fibre for dispersion management and only 7 in-line fluoride-based EDFAs were used. The fluoride-based EDFAs provide a relatively flat gain over the wavelength range 1530–1560 nm, allowing large channel spacings, compatible with field requirements. Here, 16 channels separated by 200 GHz and spanning 24 nm (between 1536.61 and 1560.61 nm) were used. The receiver sensitivities at 10^{-10} BER for the 16 channels show a total excursion of 3.1 dB.

Node cascadability experiments are also carried out aiming to assess the feasibility of a meshed Pan-European network. The OPEN project carried out laboratory experiments with the cross-connect architecture described in the previous section, with the use of the advanced technological components provided by the project. In particular, a nearly penalty-free cascade of three sub-equipped optical cross-connects with a total transmission distance of 1000 km on standard G.652 fibre was demonstrated, thanks to the reshaping properties of the all-optical interferometric wavelength converter [2]. The used cross-connects were representative for 4 x 4 size cross-connects, with 8 wavelengths per port and 2.5 Gbit/s per wavelength.

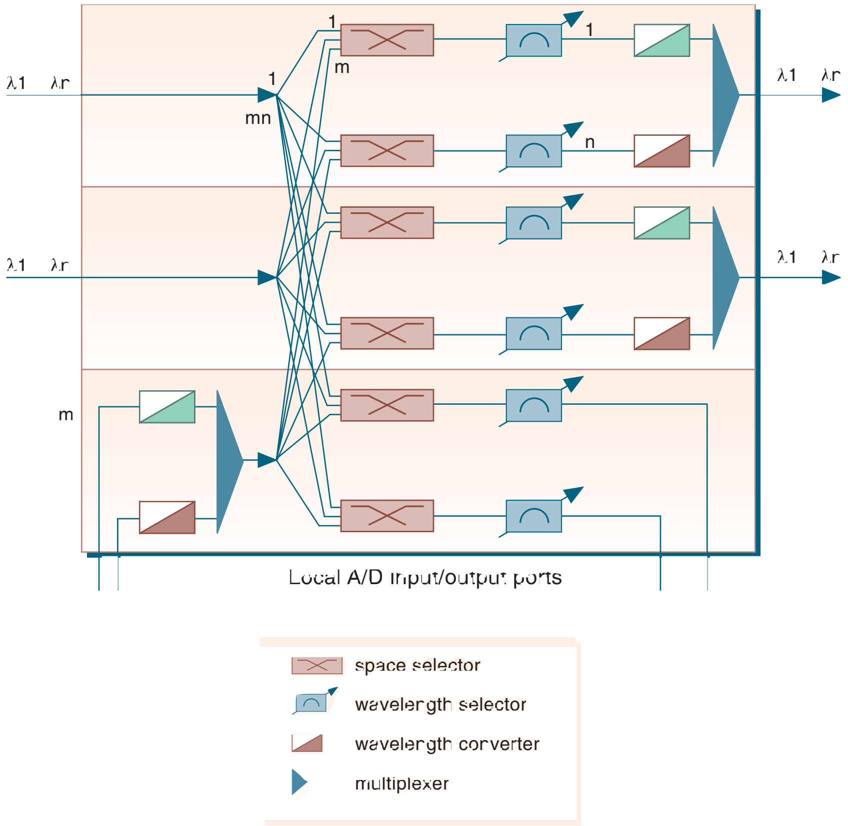


Figure 3 OPEN MOXC architecture

4 Field trials – the role of Telenor

OPEN will carry out two demonstrations in two different cross-border field trials, the first one between Norway and Denmark, and the second one between France and Belgium. The schematic layout of the trials is given in Figure 4. STM-16 signals will be transmitted and routed, using 5 channels (4 working channels and one protection channel), 400 GHz apart. Under the constraint of a limited infrastructure length, the choice was made to implement a node physical connectivity of 3 or 4 with unidirectional links rather than using fully bi-directional links. The OPEN project will demonstrate the powerfulness of its approach in using the same cross-connect, amplifiers and chromatic dispersion compensation units in the two different field trial environments. In both field trials long distance WDM transmission and routing experiments, as well as network restoration demonstration, will be performed. Configuration, performance, alarms and protection management will be demonstrated. Some important demonstrations are planned: in the N-DK field trial, critical issues are four wave mixing (FWM) effects and very long repeater spacings in the two hybrid G.652/G.653 links bridging Arendal to Thisted either through Kristiansand or through Hjørring; in the F-B field trial, the span of WDM transmission and routing experiments can reach more than 1500 km with a cascade of four wavelength translating cross-connects in addition to other types of nodes (wavelength routing cross-connect, crossbar-switch). Fast restoration is another issue for this field trial.

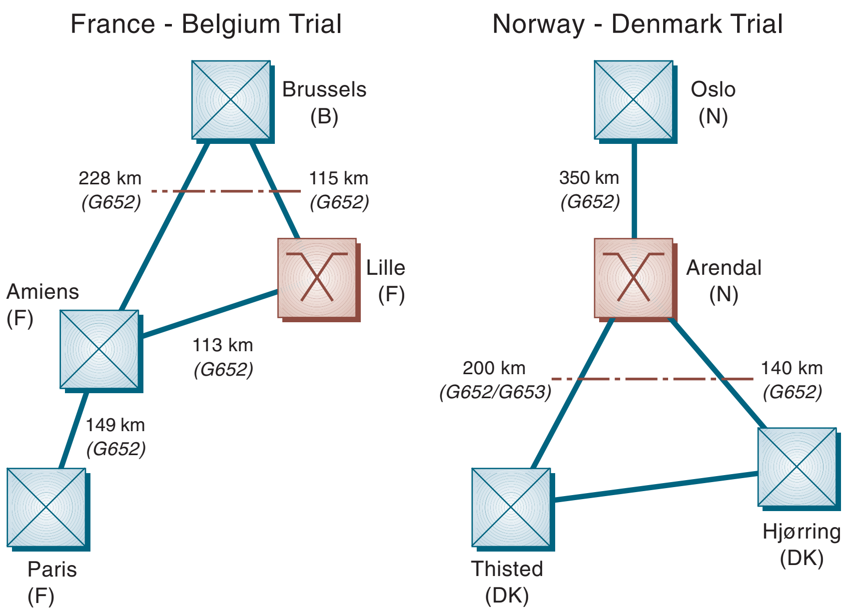


Figure 4 Schematic representation of the OPEN field trials

The role of Telenor is in the network studies and primarily in the planning and facilitation of the Norway-Denmark field trial. Prior to the field trial, a comprehensive measurement programme has been carried out to determine the fibre parameters in terms of attenuation, chromatic and polarisation mode dispersion. In particular, much attention has been paid to the study of the effect of FWM in the dispersion shifted fibres in submarine cables. Determination of the zero dispersion wavelength has been an important input parameter to the analysis of this subject. The measurement data from the fibre characterisation have been fed into transmission models provided by several partners to simulate the field trial network performance and optimise network design, such as amplifier spacing and output power levels. Significant contributions have been made by SINTEF, the sub-contractor of Telenor R&D, in the study of the Pan-European network architecture, transmission limitations and analysis of field trial network performance.

Figure 5 shows the network elements and fibre infrastructure of the N-DK trial network. The network consists of a transmit node in Oslo, a central node with the optical cross-connect and add-drop facilities located in Arendal, a transit node in Hjørring with a channel drop option, and finally, the terminal node in Thisted. The OXC and the line amplifiers featuring state-of-the-art technology are supplied by Alcatel Alsthom Recherche,

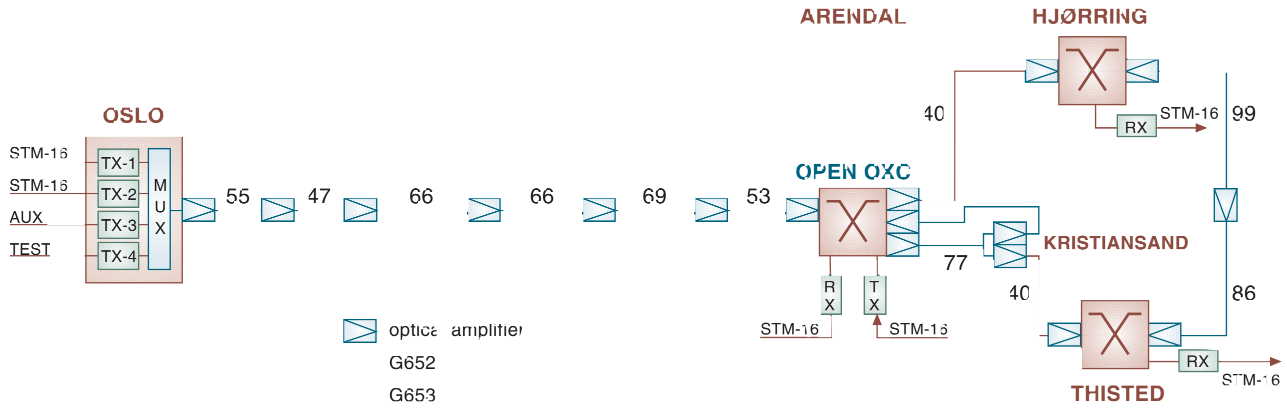


Figure 5 Network elements and fibre infrastructure in the N-DK field trial

while the terminal equipment is provided by the telecom operators. An important feature of the trial network is the interface to the existing national SDH-network, which is implemented by means of STM-16 interface transponders developed by Telenor R&D. All planning activities and operations related to the N-DK field trial are carried out in close co-operation with TeleDanmark and SINTEF, and strongly supported by Alcatel Alsthom Recherche and Alcatel Telecom Norway. The N-DK field trial installations, evaluation programme and demonstrations are going to take place in the period October-December 1997. The OXC and the line amplifiers, supplied by Alcatel, are subsequently shipped to the sites of the F-B field trial for complementary tests.

Both field trials are aiming at the validation of the future all-optical Pan-European network and proving the new technology in real operating environments.

Acknowledgements

The author wishes to acknowledge the key managers of the OPEN Consortium, M. Chbat, A. Leclert, E. Jones, B. Mikkelsen, N. Flaarønning, T. van den Broek, M. Pouleo, H. Melchior, P. Demeester, P. Gødsvang, A. Jourdan, and D.R. Hjelme for contributing to the source material on which this article is based.

Torodd Olsen is Research Scientist at Telenor Research & Development. He has been employed by Norwegian Telecom since 1976, apart from a period with Tandberg AS from 1980 to 1986, when he started at Telenor R&D working with optical high speed and coherent communications systems. Since 1995 he has been Project Manager in optical communications, and is currently manager of the Optical Networking group.

e-mail: torodd.olsen@fou.telenor.no

References

1 Artigaud, S et al. Transmission of 16x10 Gbit/s channels spanning 24 nm over 531 km of conventional single-mode fibre using 7 in-line fluoride-based EDFAs. *Conference on Optical Fibre Communications (OFC'96),* San Jose, CA, February 26 - March 1, 1996, post-deadline paper PD27.

2 Jourdan, A et al. Experimental demonstration of penalty-free optical crossconnect cascade with multiwavelength transmission over 1000 km standard fibre. *Conference on Optical Fiber Communications (OFC '97),* Dallas, Feb. 97, paper TuR3.

3 Chbat, M W et al. WDM for transmission and routing in wide-scale networks : the Optical Pan-European Network approach (OPEN). *NOC'96,* vol. 1, Broadband Superhighway, 182–186, Heidelberg, June 1996.

4 Chbat, M W et al. Pan-European Optical Networking using Wavelength Division Multiplexing. *IEEE Communications Magazine,* 35, (4), 1997.

List of participants in the OPEN Consortium

F	Alcatel Alsthom Recherche
N	Alcatel Telecom Norway
B	Belgacom SA
I	Centro Studi e Laboratori Telecomunicazioni – CSELT
CH	Eidgenössische Technische Hochschule Zürich – ETH-Z
F	France Telecom – CNET
B	Interuniversity Micro-Electronics Centre IMEC
N	SINTEF
DK	Technical University of Denmark
DK	TeleDanmark
N	Telenor
UK	University of Essex.

MoMuSys

ROBERT DANIELSEN

1 Introduction

MoMuSys (Mobile Multimedia Systems) is project number AC098 in the ACTS (Advanced Communication Technology and Services) programme, funded up to 50 % from the European Commission. The rest is funded by the partners involved. The project started on 1 September 1995, and will run until autumn 1998. The project leader (prime) is Dirk Lappe from Bosch. The MoMuSys project has already produced good results, and is regarded by the Commission as a very successful project.

1.1 Project purpose

From the official project description:

> *"The main objective of the MoMuSys proposal is the development of mobile multimedia technology for real use, flexible enough to achieve all functionalities that will be required in future multimedia applications. This will be performed by developing tools, algorithms and profiles which will be validated in usage trials on real time hardware. Close collaboration with the MPEG-4 standardisation effort will ensure the international relevance of the work."*

In addition to this widely defined goal, one of the main reasons for the EU to support the project is that it boosts the European level of research and development and strengthens the industry in this very important area.

Developing new software-based solutions for multimedia communication is a very important part of the project. This work fits very well with the rapidly increasing use of the Internet and of Multimedia applications in general. The fact that MoMuSys is very active in the ISO/MPEG-4 development means that this part of the work is linked to standardisation and is therefore open for use outside the MoMuSys project. This may prove to be very beneficial to MoMuSys, the industry and users in general. The technology developed will not be hidden, but will be part of an open standard.

1.2 EU ACTS initiative

MoMuSys is intended to be a European initiative for pushing technical development and standardisation for multimedia. The ACTS programme provides the basis for forming a European collaboration platform to influence international standardisation and to push the technical development in Europe.

From the ACTS sub-domain on Multimedia Content Manipulation and Management (MCM) newsletter we read he following:

> *"If the primary objective of the ACTS programme is to conduct R&D in the context of trials, it will also be strongly supporting leading-edge research aimed at developing the tools for the future multimedia landscape. Within ACTS, the purpose of the MCM domain will be to develop audio-visual signal representation, manipulation and management tools able to support the creation of advanced digital multimedia services. More specifically, the activities of the group will cover multimedia signal analysis and coding with an emphasis on contributions to MPEG-4. It will also include synthetic and natural hybrid coding (MPEG-4 SNHC group), syntax (MPEG-4 MSDL group), 3D analysis, modelling, texture mapping and VR technology. Applications will range from 3D tele-presence to mobile multimedia communications, programme production and generation of distributed virtual environments."* (Eric Badiqué, Commission Project Officer.)

The MoMuSys project covers parts of all the areas mentioned in the above paragraph. Through the close work with MPEG-4 we are contributing to the solution of relevant problems in all these areas.

1.3 Area of work

The boundaries between telecommunications, computers and TV/film industries are blurring. MoMuSys will address all three areas through the field trials. The multimedia coding technology will be developed and promoted through the MPEG standardisation work.

In addition to developing necessary technology for the new functionalities addressed by MPEG, MoMuSys will work on developing the necessary network-related technology, particularly for mobile networks.

To validate the technical solutions developed, MoMuSys will carry out three specific user trials, where the latest technology will be deployed, and assessments and feedback from the users will be reported.

1.4 Project partners

The project partners are as follows:

Organisation	Short name	Country
Robert Bosch GmbH	BOSCH	Germany
Armines	CMM	France
France Telecom CNET	CNET	France
Centro Studi E Laboratori Telecommunicazioni SPA	CSELT	Italy
Teltec Ireland DCU	TELTEC	Ireland
Deutsche Telecom AG FTZ	DTAG	Germany
Fraunhofer Gesellschaft zur Förderung der Angewandten Forschung E.V	FHG	Germany
Heinrich Hertz Institut für Nachrichtentechnik Berlin GmbH	HHI	Germany
Instituto Suprior Tecnico	IST	Portugal
Royal PTT Nederland KPN Research	KPN	Netherlands
Laboratoires d'Electronique Philips	LEP	France
Queen Mary and Westfield College	QMW	UK
Siemens AG	SAG	Germany
Telefonica Investigacion y Desarrollo	TID	Spain
Universität Hannover	TUH	Germany

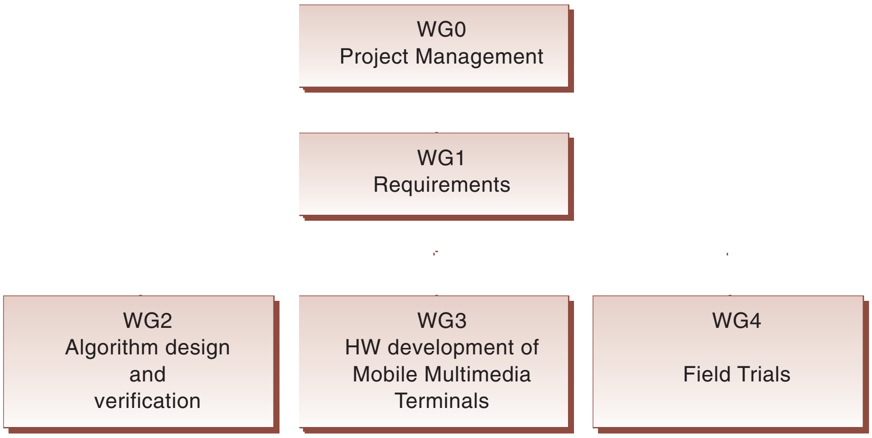


Figure 1 Project structure of the MoMuSys project

Universidad Politecnica de Catalunya	UPC	Spain
Universidad Politecnica de Madrid	UPM	Spain
Telenor AS	Telenor	Norway
VTT Information Technology	VTT	Finland
Centre Commun d'Etudes de Telediffusion et Tele-communications	CCETT	France

The project structure is shown in Figure 1.

1.5 Workgroups and workpackages

The project is divided into workgroups and workpackages as follows:

WG0	Project management	Dirk Lappe,	Bosch
WG1	Requirements	Karlheinz Brandenburg,	FhG
WP1.1	Internal requirements		
WP1.2	External requirements		
WG2	Algorithm development	Liam Ward,	Teltec
WP2.1	Integration		
WP2.2	Syntax and Control		
WP2.3	Network Adaptation		
WP2.4	Video Algorithms		
WP2.5	Audio Algorithms		
WP2.6	SNHC		
WP2.7	Analysis and Coding at higher bitrates		
WP2.8	Audio channel coding, error robustness and error concealment for audio		
WG3	Development of Mobile Multimedia Terminals	Alan Pearmain,	QMW
WG3.1	Multimedia Terminal Hardware		
WG3.2	Multimedia Terminal Software Algorithms		
WG3.3	Integration and Test		
WG4	Field Trials	Väinö Typpi,	VTT
WG4.1	Field Trials with End User Groups		
WG4.2	End User Systems Design and Implementation		
WG4.3	MoMuSys User Group Trials		

2 Technology areas

MoMuSys is working to develop new ways of coding and communicating multimedia information. To achieve this, MoMuSys is working in several fields of technology. To be able to develop full solutions that can be tested on non-technical users it is necessary to go further than producing research results. We need to develop actual solutions in software and hardware, after first developing the concepts and doing the system designs.

2.1 Mobile telecommunication

MoMuSys is concerned with *mobile* multimedia communication. The project will consider several available technologies for mobile communications, networks for both high bitrates and low bitrates.

The actual networks to be included in the trials are:

ISDN over DECT	User trial with ordinary users in a test region in Finland
GSM and PSTN	Very low bitrate testing
Wireless LAN	High bitrate testing
UMTS	3rd generation mobile communication.

Other networks may come into consideration during the project.

Common to all these networks in the first two years is the fact that we will be running the communication on the TCP/IP protocol. In the third year we plan to use unprotected mobile channels and add our own error protection layers.

2.2 Signal processing

Much of the activity is related to digital signal processing. This is mainly done in the audio and video groups, which are developing new processing techniques for audio and video. We also have an SNHC group, which works with Face and Body Animation.

2.2.1 Audio processing

The MoMuSys Audio group is led by Luca Cellario from CSELT.

In the audio group the work on developing new techniques is done through active participation in the MPEG audio group. The MoMuSys project is a driving force in this work, especially through the active role of Bernhard Grill from FhG.

The MPEG-4 audio standard is in its current draft state a very broad standard. It covers a wide range of bitrates, and also a wide range of aural bandwidth. The current method consists of several algorithms:

- AAC (Advanced Audio Coding). This is the high level algorithm, covering bitrates from approx. 8 to 256 kbps, with the normal bitrate at 64 kbps. This algorithm is a time/frequency method (transform method) with additional vector quantization tools.
- Broadband CELP (Code Excited Linear Prediction). This method is a well known algorithm, and is here used for coding 7 kHz audio at bitrates mainly in the area 16 – 32 kbps.
- Narrowband CELP. This is similar to the above method, but for coding speech at 3.1 kHz. The bitrates are normally around 8 – 32 kbps.
- Parametric codec. This is an effective speech coder, with low bandwidth and low bitrate. The bitrates are typically from 2 – 8 kbps.

This list of parameters for the algorithms quickly gets outdated, since the proponents of the different algorithms are adding features and extending both the bitrates and the aural bandwidth covered. One idea with this system of several algorithms is that it will be possible to change algorithm “on the fly”, that is, when the bitrate changes, the audio coding method changes adaptively. In modern networks this is a very useful feature.

The audio group is also working to develop a very exciting aspect of MPEG-4, namely the coding of individual objects. Audio objects may be coded separately, either after being separated by analysis or by separate recording. The audio objects may then be transmitted with different quality, bitrate or by different algorithms. This allows for new ways of interacting with audio material.

2.2.2 Video processing

The MoMuSys Video group is led by Ferran Marques from UPC.

Like audio the video coding activity in MoMuSys is closely linked with MPEG. The video coding method developed in MPEG was implemented in MoMuSys from a very early stage, and the MoMuSys implementation is one of the two official reference implementations of MPEG-4 video. This is an advantage for MoMuSys, since it has resulted in a very strong position in MPEG. It also means that there is a burden on MoMuSys to always keep the software up-to-date with the latest changes made at MPEG meetings. However, new proposals are often being developed by other companies on the MoMuSys software, and after acceptance in MPEG, the software is donated to MPEG and integrated by MoMuSys to become a part of the official reference software.

The texture coding is based on the ITU H.263 standard. One may say that this is the core of the method. Nevertheless, the method has gone through a number of changes, and a very important new dimension has been added, namely shape coding. The core software based on H.263 was developed by Telenor earlier as part of the standardisation work for H.263, and donated to MoMuSys as background code. This software has been very effective in promoting Telenor both in ITU, MPEG and in MoMuSys.

Shape coding implies that the video scene is actually automatically analysed to partition the scene into separate video objects. This process is called segmentation, and is very computationally intensive. MPEG will not standardise a segmentation method, only the coded representation of the shapes. This means that it is very important to develop good tools for segmentation, since they will not be provided through MPEG. There will probably be an example implementation in the informative part of the standard. The MoMuSys partners are working actively on developing such algorithms for segmentation. Many of the partners are also involved in the European project COST211ter, which is concerned with developing good techniques for segmentation.

After segmentation, MPEG standardises the coding of the shape information, the texture and the motion information. The coding method for the shape information was developed by Noel Brady from MoMuSys (Teltec) and Frank Bossen from EPFL.

The video objects may actually be treated like individual streams, and such streams may come from different locations and be composed into the same scene. More about scene composition below.

2.3 Multimedia applications

The application areas for MPEG-4 may be very diverse. From MPEG-1 and MPEG-2 we know that these standards were targeted at special application areas. MPEG-1 was for digital storage (CD-ROMs) and MPEG-2 was mainly aimed at TV-quality broadcasting, although the video part was also standardised by the ITU, and is thus suitable for two-way communication (using the low delay tools).

In contrast to this, MPEG-4 is aiming at a much broader range of application areas. At the low end MPEG-4 covers the very low bitrates and may be used for mobile audio-visual communication, also two-way communication. For the high end MPEG-4 has added the same interlaced tools as MPEG-2 and may be used for the same broadcast type applications. Recently, MPEG has also been adding a separate still image algorithm.

In the MoMuSys project we are mainly concerned with mobile communication, and therefore the work has been focused on the low bitrates. Still, MoMuSys has added an extension to the project for working on the higher bitrates.

One of the main features that distinguishes the MPEG-4 from other standards is the object based coding of audio-visual signals.

2.4 System communication and composition

The MoMuSys Systems group is led by Robert Danielsen from Telenor.

Apart from the coding of audio and video signals (both natural and synthetic), the system design of MPEG-4 is a very exciting development. The MPEG-4 System specification will probably prove to be the most important part of MPEG. Even here MoMuSys has assumed a central position. Olivier Avaro from MoMuSys (CNET) is the leader of the MPEG Systems group, and MoMuSys has contributed considerably to the common reference software development in the group.

The main features of the system is reflected in the architecture, where elementary streams are coming from a demultiplexer into an MPEG-4 decoder system. From the Binary Format for Scene Description (BIFS) the player builds up a scene graph, and an executive module ensures that decoding, rendering and presentation is being done in a synchronised way. The compositor module builds up and controls the scene graph, and takes care that all rendering and presentation is done. Figure 2 illustrates the architecture.

Figure 2 shows how the elementary streams from the multiplex are connected to nodes in the scene graph. These nodes are actual decoders, and the output is treated by the compositor, which carries out the rendering of the individual sources and presents the data to the user. The composition information which describes the placement of audio, video and other sources in the scene is transmitted in a compact description (BIFS).

The system work is closely related to MPEG-4 (development group "im1"). The successful 3D based player made by MoMuSys will be merged with a 2D player made by CSELT, and will be the official reference software for MPEG-4 Systems. The 3D compositor was developed by Karl Anders Øygard from Telenor. In the period July – November four people from MoMuSys have direct responsibility for developing parts of the MPEG-4 system software (3D/2D Compositor, Multiplex, BIFS, Audio composition).

The MoMuSys Systems group will build further on this platform and add parts that are necessary for our field trials. Some of these parts may in turn be contributed to MPEG.

2.5 Combining natural and synthetic data

One of the new areas in the MPEG standardisation work is covered in the SNHC group (Synthetic Natural Hybrid Coding), where synthetic data is combined with natural data. The background for this work is that synthetic data are not very efficiently coded with coding methods developed for natural data and vice versa.

Some work in this area is being carried out in MoMuSys, in workpackage 2SNHC. However, the work has been mostly related to the system level composition, using FAP (Facial Animation Parameters) to build animated faces into scenes. This kind of work is also done in workpackage 2.1 (Integration).

In the MPEG system software development, Miralab from Geneva has made very good contributions to Face and Body Animation (FBA). This software is currently being used in the 3D compositor developed by Telenor for MoMuSys. This was demonstrated at the 40th MPEG-meeting in Stockholm 21–25 July 1997.

2.6 Extensions to higher bitrate networks

The project has been extended to take into account the scaleable coding for higher bitrate networks. The work will be focused on video coding. Both scaleable texture coding and scaleable shape coding will be worked on, and contributions will be made to MPEG.

This work is to be done in WP2.7, and will also cover work on algorithmic parallelisation of coding tools.

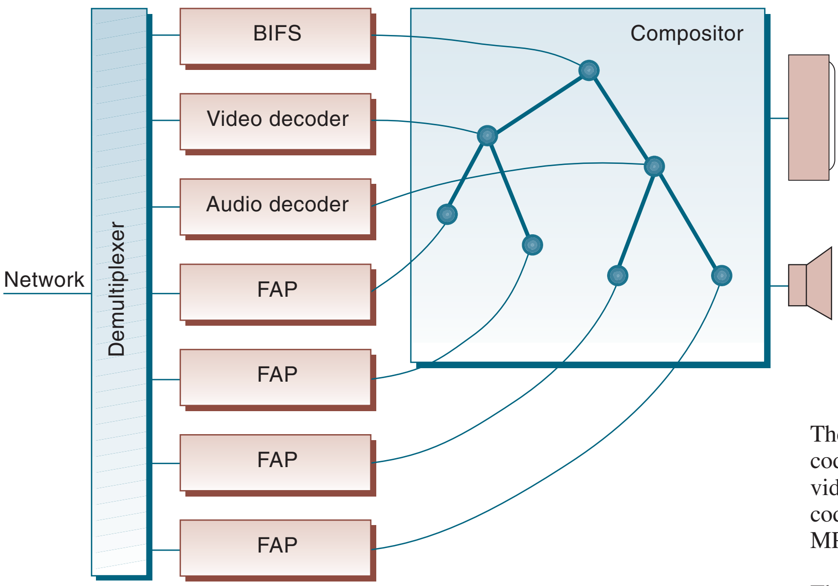


Figure 2 System architecture

2.7 Hardware-related development

This work is carried out in WG3, led by Alan Pearmain from QMW.

This group is responsible for specifying the appropriate hardware terminal to be used for running the software solutions from WG2 in the MoMuSys field trials. In the first phase terminal we are using a DSP-board in a portable Pentium based machine. For phase 2 we will include an additional DSP card for increased processing power.

Most of the partners will purchase identical systems, to follow the specification from WG3. We will then have a base of these systems, which can interoperate in various networks. The terminal will also be used in other ACTS-projects, for Telenor we note the projects CRABS and SINUS.

2.8 Field trials

The MoMuSys field trials will be carried out in three phases. The trials will be done with "End users" and "Expert users".

The work with end users is conducted within the frames of two scenarios, "Traveller's AV Information Services for Mobile Users" and "Mobile Remote Security Surveillance". The first scenario is very mass market oriented in nature and, accordingly, is targeted to layman users. It mainly deals with AV database access. The second one is targeted to professional users in the security service field dealing principally with real-time AV information transfer.

3 Strong links to international standardisation

3.1 MPEG

MPEG (Moving Pictures Expert Group) is a nickname for an ISO subgroup:

ISO/IEC JTC1/SC29/WG11

ISO	International Standardisation Organisation
IEC	International Electrotechnical Commission
JTC1	Joint Technical Committee 1 (Information Technology)
SC29	Subcommittee 29 (Coding of audio, picture, multimedia and hypermedia information)
WG11	Coding of moving pictures and audio

The connection to the work in MPEG is very strong in MoMuSys. The work in MPEG is very important for the standardised development of coding methods for multimedia signals. The Commission also appreciates this, and is very keen to contribute to the MPEG work.

3.2 ITU

The project is also contributing to work in the ITU (International Telecommunication Union) in the fields of video and audio coding, and also in the field of multiplexing and mobile communication.

3.3 Other external bodies

The project is actively participating in other standardisation bodies and consortia. People in the project also contribute to conferences and workshops in the area of coding and communication. All this contributes to the successful promotion of the MoMuSys work.

4 Results achieved to date

The MoMuSys project has been a success and has contributed significantly to the development of standard solutions for the coding and transmission of multimedia information.

4.1 Software

Software has been contributed to MPEG for Video coding, Audio coding and Systems. This software will be donated to MPEG, and will become part of the actual MPEG-4 standard. This will benefit the industry and all users of the standard. MPEG is also considering interoperability with other standards, and may prove to become a very flexible and useful standard.

4.2 The MoMuSys terminal

The first phase MoMuSys terminal has been specified, and software has been developed and optimized for the chosen platform. Most of the partners have acquired the terminal and will use it for testing and evaluation, and for demonstrations.

Work is underway for designing the next version of the terminal, which will be ready in January 1998.

5 Project management in the EU

The MoMuSys project has proved itself to be a success. It consists of 20 partners, a large number for a research project. The reasons for the success are probably several:

- Enthusiasm about the field of research. The MPEG-related work is very relevant for the continued developments in Multimedia-related fields like Internet, videoconferencing, broadcasting and digital storage. We will see other applications maturing in the years to come. MoMuSys and MPEG are working to provide the foundation for further developments in these areas.
- The project management works very well, and the work is being followed up closely from the European Commission. There are frequent deliverable reports to be written, with well planned schedules.
- The work in the project is highly relevant to manufacturers, service providers and universities. We have a good mix of these in the project, and the co-operation is going surprisingly well.
- The social side of the project is also not to be forgotten. At every meeting we arrange a sports or leisure activity, which certainly also contribute to the informal and friendly atmosphere in the project.

6 Continuing the work

6.1 Each partner to his own

All the partners will get a strong foundation for further developments in the area of multimedia coding and communication. The partners are not only observing the industry and acting accordingly, we are actually shaping parts of this industry through our research work and contributions to standardisation.

6.2 Strengthening European industry

Several of the partners have internal product developments related to the MoMuSys work. The fact that MoMuSys is so active in standardisation, and is actually developing the reference implementations for the industry means that the partners are in a very good position to be first on the market with their new products.

6.3 Towards MPEG-4 extensions and MPEG-7

The discussions on extensions and of starting new projects have begun in the MoMuSys group. One of the very interesting new developments is to be seen in MPEG-7, which will be concerned with content description.

In view of the rapid deployment of the multimedia coding standards, there is an increasing amount of multimedia data existing on the Internet, in databases, with broadcasters, in live feeds, etc. To be able to find a specific sequence of video or data today one has to look at the data itself. MPEG-7 will provide a standardised solution for this, by tagging the contents with descriptions of the content. These descriptions may then be accessed by agents or search tools to find the needed information.

7 Conclusions

As seen from Telenor's viewpoint, the work in the MoMuSys project is very satisfying. It provides a base for highly relevant research and development, with a 3-year contract. This is very important for keeping up a sustained effort in this field of work.

The project is very successful, receives wide acclaim, and has dedicated partners who have a genuine interest in ensuring the success. This project certainly shows that the ACTS programme works very well in this area.

Robert Danielsen is Research Scientist at Telenor R&D. He has been working with video coding in international standardisation and projects for several years. Currently, his interests are in the development of standardised solutions for multimedia coding in MPEG-4, and the commercial exploitation of this work in Telenor.

email address: rid@nta.no

The SINUS project, Satellite communication in the UMTS

BY VENDELA PAXAL

The SINUS project, Satellite Integration into Networks for UMTS Services, is an ACTS project (AC212) that started on 1 July 1996, and has a planned lifetime of about two and a half years, with a total human resource allocation of 666 Man-Months (MM). The project's consortium consists of nine partners, which are listed in Table 1 with the respective MM allocations. The main objective is to define techniques for integration of the satellite component into the UMTS (Universal Mobile Telecommunication System), with special focus on the interworking with the terrestrial component. This interworking contains critical aspects such as intersegment handover, call handling and mobility management. In SINUS, several of these aspects will be studied and the developed solutions will be demonstrated. The particular features of the mobile satellite systems, including Doppler frequency shifts, important propagation delay, low transmit power budget and severe radiowave propagation environment, necessitate a careful design of the modem and the air interface. Hence, an equivalently important goal of the SINUS project is to find solutions to these problems, and to define reference models in order to propose a common European platform for meeting the challenge of such system constraints. Further, participation in standardisation bodies is viewed as part of the outcome of the SINUS project.

Three major phases of the SINUS project have been defined. The general system study and the SINUS test-bed specification are comprised in the first part, development of the testbed constitutes the second, and finally the demonstrations will be performed in the third part. In the SINUS test-bed there will be a mobile terminal connected to a fixed earth station through a satellite channel simulator. The fixed earth station provides access to a core B-ISDN network represented by the National Host ItalHost in Torino, and a transparent terrestrial emulator will ensure the signalling link necessary to demonstrate intersegment handover algorithms. The SINUS project is today about to finish the first phase, and is thus entering the second. Co-operation with other ACTS projects has been initiated, such as co-operation concerning the radio access interface with FRAMES, similar system architecture descriptions and protocol definitions with RAINBOW, and finally an agreement for using the multimedia terminal developed by MOMUSYS as an application for the SINUS demonstrations.

Table 1 SINUS consortium partners and resource allocations

Status	Partner	Country	MM allocation
Partner 1, prime	Alcatel Espace	F	150
Associated partner, partner 1	France Telecom CNET	F	30
Partner 2	CSELT	I	58
Partner 3	Matra Marconi Space	UK	24
Partner 4	Bradford University	UK	56
Partner 5	Surrey University	UK	62
Partner 6	Fundacion Airtel Movil	E	28
Partner 7	Telenor R&D	N	44
Partner 8	VTT Electronics	SF	73
Sub-contractor, partner 8	Elektrobit Ltd	SF	11
Partner 9	SINTEF Telecom & Informatics	N	130

1 Introduction

The goal of the SINUS demonstrations is to show the feasibility of an efficient integration of the satellite component into the terrestrial cellular UMTS network. As a consequence, the problems related to this integration will be identified, studied and adequate solutions will be proposed. The terrestrial approach to UMTS is different from what is natural from a satellite oriented point of view, which is due to the different types of radio propagation environment, with their restrictions and possibilities. Concerning the integration of one component into the other, it is a goal to achieve as high similarity between the technical solutions in the two systems as possible. Hence, this becomes an extra challenge to the system developer, since solutions found for the terrestrial system are not always optimal in a satellite system and vice versa. In Europe, the development of a terrestrial solution is in advance on a satellite solution, which means that the development of S-UMTS will have to follow the terrestrial solutions to a great extent. This will not be possible in all cases, thus requiring some dual mode operation blocks in terminals designed for operation in both satellite and terrestrial UMTS networks.

One of the consequences of the integration of the S-UMTS into the T-UMTS, is the adoption of the GRAN (Generic Radio Access Network) principle for describing the system architecture in the SINUS project. This GRAN principle has been described in [1] and [2], and for SINUS, the main interest in this concept at the actual stage of the project, is the division of the architecture into radio access dependent and radio access independent parts. In this manner the architecture description approach used in the SINUS project is compatible with the ITU recommendations and the other European projects working on different aspects of the UMTS. The protocol layer description and the application of the GRAN principle to this description are developed in section 3 dedicated to the SINUS test-bed presentation. We will in section 2 concentrate on the satellite system particularities and the critical aspects of the mobile satellite communication and the integration into the terrestrial component.

2 SINUS demonstrations

It was natural to start the study in SINUS by identifying the particular problems encountered on a mobile satellite link, and then turn towards the difficulties of an integration into the T-UMTS. The radio propagation environment for mobile satellite communication is characterised by rapid variations due to the movement of both the satellite and the terminal. The evolution of these variations depends on the satellite constellation earth orbit and the environment in which the terminal moves. In SINUS the LEO, MEO and GEO constellations have been selected for

Table 2 Characteristics for candidate satellite constellations

Characteristics	LEO	MEO	GEO	Units/ Notes
Orbit height	700–4000	10000–25000	36000	km
Propagation delay	5-25	80-120	280	ms
Satellite handover during call	Every 10 min.	Every 2 hours	Unlikely	Typical values
Beam handover during call	Every min.	Every 10 min.	May occur	Typical values
Nb of satellites, near global coverage	>48	10-15	3	

study (the HEO constellation has been omitted because of coverage limitations and important delay). The characteristics of the different orbits are resumed in Table 2, where we also find an indication of the passing time. The environments in which the terminals should function are; rural, suburban, and urban, the indoor situation will be studied during experimentation but we know that this type of communication may prove difficult and even impossible. The terminal types will englobe the handheld, the vehicular, the transportable and the fixed installation. It is clear that even within the mobile satellite communication area defined here, huge variations will be found between scenarios made out of different associations of elements from the three categories satellite constellation, radio propagation environment and terminal type.

The impairments met in a satellite channel are principally the shadowing or blockage, the Doppler carrier frequency shift due to rapid satellite motion, high transmission delays for some constellation types, interference (from other systems, other channels, other users ...) and the low power budgets. Satellite communication has always suffered from restrictive use of available power due to the satellite component power supply. In mobile satellite communications, we also meet power restrictions due to the use of for instance a handheld terminal (not only due to battery limitations but also for medical reasons.) These power restrictions will amplify the shadowing and blocking problem. The Doppler carrier frequency shift may be compensated to some extent, but will always necessitate a robust carrier frequency tracking at the receiver. The transmission delay will limit real-time, interactive applications. Signal processing algorithms depending on information from the remote terminal, such as for instance power control, will also prove sensitive to transmission delay. Different types of interference constitute a complex problem, and much effort has been made over the past years in order to determine the best access method to meet these problems. No unique answer seems to have been given, since all methods have their particular advantages and disadvantages and they are dependent on system parameters such as transmission delay, required bandwidth and carrier frequency, number of users and traffic burstiness to mention only a few. This is why many projects propose the use of hybrid schemes, thus meeting both different types of impairments in a quasi-optimal way, and also different types of system designs for multimode connections, at the expense of a higher complexity. This problem has been of major concern to the SINUS consortium, and much effort has been put into the co-ordination with other projects, in particular those describing the terrestrial solution, in order to achieve as much commonality as possible on this issue. The final choice in SINUS will be oriented towards a CDMA solution.

The integration of the satellite component into the terrestrial network requires the development of powerful network management algorithms such as mobility management, call set-up and release and handover. The desire to keep the terminal design at a reasonable complexity level imposes that most of the network management evaluation, decision and execution are performed by the network and not the terminal. In SINUS, the study of these functionalities has been one of the major tasks in the first phase of the project. In the project definition phase, it was decided to demonstrate the call set-up and release and the major aspects of handover, but not the mobility management due to limitations of the test-bed set-up (limited number of transmission channels and terminals). The mobility management has however been studied in the first phase of the project, and the functionality covers both the identification of the geographical location and the identification of the network to which the user is connected. Procedures such as registration/deregistration and location update will be initiated by the mobile terminal, whereas the connection set-up procedure will be handled by the network. The call set-up and release functionalities are grouped into two main parts, the session set-up/release and the call set-up/release. The session set-up comprises four actions which are the service provider selection, the access control verification, the set-up of the terminal connection and finally of the user connection. For the call set-up, the called user is located, the service profile of the called user is verified together with the domain and terminal capabilities, then follows the paging, the terminating session and the bearer connection. The paging informs the called user that there is an incoming call, and if the called terminal is mobile, the network initiates a terminating session before the bearer service is connected.

Two types of handover are required for the integration of a mobile satellite communication system into the terrestrial; inter- and intra-segment handover. The inter-segment handover concerns the handover from the satellite component to the terrestrial or vice versa, whereas the intra-segment handover is related to the switching between two spot-beams or satellites during a communication. The handover strategies selected depend on the amount of processing considered reasonable in the mobile terminal together with the handover performance. Some of the requirements in UMTS are that the handover should in principle be seamless in order not to degrade the QoS, the number of handovers should be kept to a minimum and handover between different operators and in different environments should be possible. The compromise between these requirements and the implementation complexity considerations leads to the selection of a backward strategy associated with a hard mobile assisted handover with signalling diversity. Hard handover will lead to an interruption in the connection, but the duration of this interruption will be reduced by the signalling diversity technique.

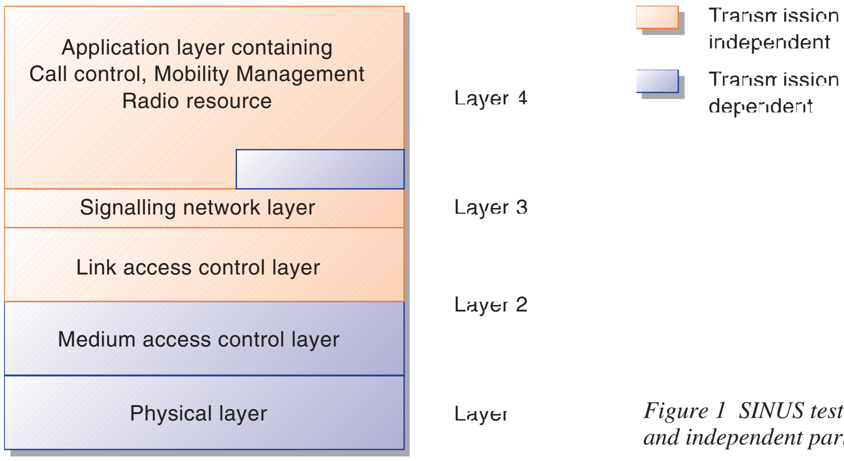


Figure 1 SINUS test-bed protocol layers with radio dependent and independent parts

The signalling diversity implies that a connection is established through the new link before the hard handover is performed. The network will be the "brain" behind the initiation and execution of the handover algorithms, only assisted by the mobile terminal in order to reduce the complexity in this unit.

The type of services selected for the SINUS demonstrations are offered through a multimedia application. The choice was motivated by the possibility of demonstrating different services such as speech, video and data transfer, or a combination of any of these, all from the same terminal. The environment and type of satellite constellation are expected to have varying impacts on the QoS of the different types of services.

3 SINUS test-bed description

As mentioned in section 1, the SINUS test-bed architecture is based on the GRAN (Generic Radio Access Network) principle. This implies that the radio interface will be compatible with a variety of core networks (GSM, N/B-ISDN, ...), all included in a UMTS Generic Core Network. Figure 1 represents the protocol layers defined for the SINUS test-bed. The four protocol layers are the application layer, the signalling network layer, the access control layer and the physical layer. One of the main purposes of the GRAN approach is to identify the radio dependent and radio independent parts of the architecture. This permits a separation of the parts of the terminal which are interchangeable for any type of access network with those having to be modified according to the radio access network the terminal will be connected to, and thus making the terminal design modular and as universal as possible. In Figure 1 we see that the radio dependent parts are concentrated on the lower network layers (shadowed in the figure). The physical layer is obviously radio dependent, and so is part of the access control layer. Layer two has been divided into two sub-layers, and the frontier of the radio dependency passes between those sub-layers. The sub-layer called medium access control (MAC) is dependent on the radio access, whereas the link access control (LAC) is radio independent. The MAC contains functionalities such as radio resource management and link quality control, the LAC is concerned with for instance the logical link call set-up and release functionalities. The network layer is considered independent of the radio access, so is the application layer with the exception of some radio resource management functionalities which are present also at this level.

Figure 2 shows the test-bed for the SINUS demonstrations. This test-bed consists of four main parts which are the mobile terminal (MT), the fixed earth station (FES), the satellite channel emulator and the terrestrial channel simulator. The MT and the FES contain many equivalent entities, such as the emulator interface (which will be implemented in IF for simplicity) the modem and the baseline. Roughly speaking, the interface and the modem contain the functionality of the physical protocol layer, and the baseline those of the access control layer. On the earth station side, there is a connection to the National Host (NH), ItalHost in Torino. This NH includes an ATM-laboratory, thus necessitating a B-ISDN interface between the FES and the NH. The ATM switch is connected to a core network as well as terminal equipment for user application interfaces. The SINUS project will add an intelligent network (IN) unity, which will control the network signalling and the management from the NH. The transmission channel of main interest in the SINUS project is the satellite channel. Since the SINUS consortium will not have any access to operating mobile satellite systems for the trials, an emulator has been procured as a replacement. This emulator will contain two full duplex channels, which will be configurable before and during simulation runs, in order to represent a realistic mobile satellite channel. The presence of two channels will also enable simulation of configurations including two beams or two satellites, essential for intra-segment handover execution tests. The inter-segment handover tests are the reason for the presence of a terrestrial segment simulator in the SINUS test-bed. This simulator will only convey signalling information from one side to the other in order to show the algorithmic validity of the inter-segment handover protocols. The

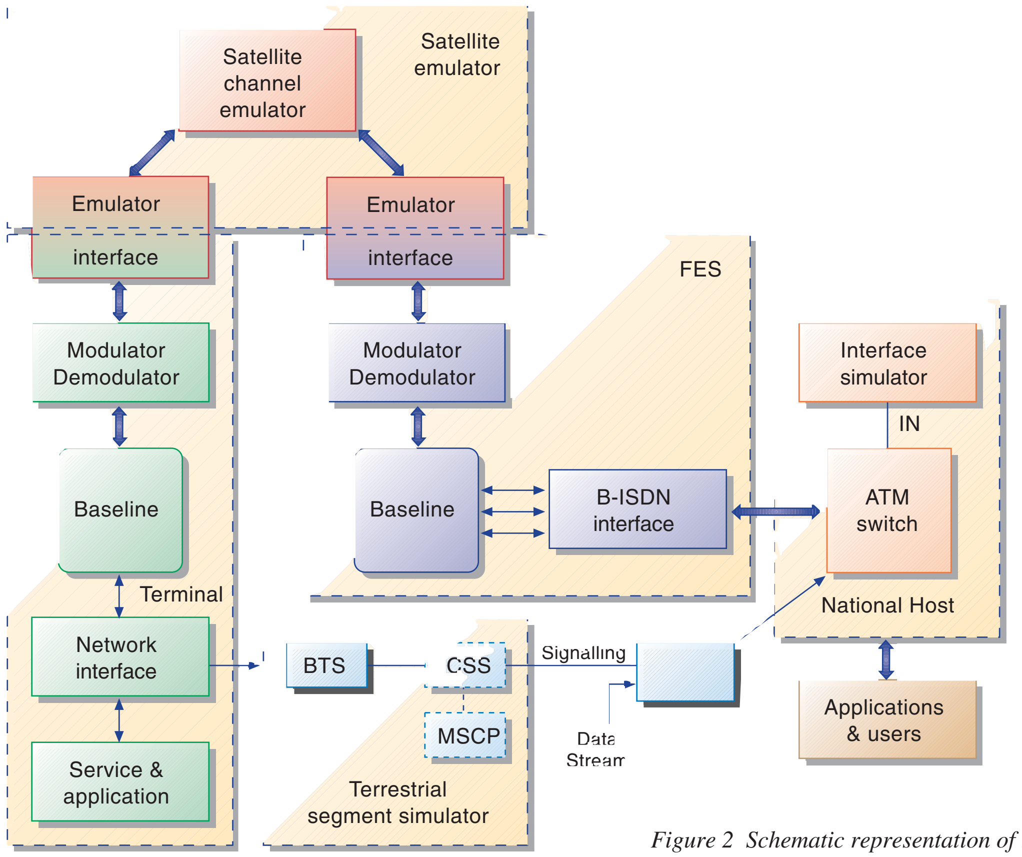


Figure 2 Schematic representation of the SINUS test-bed

terrestrial simulator will therefore be a transparent channel, i.e. it will not represent the terrestrial channel characteristics. The application data stream will not pass this terrestrial channel simulator at all, but be directly connected between the application terminals with a switch allowing the turn off and on of the data stream according to the signalling information. On the mobile terminal side, the connection with the terrestrial simulator is established through a network interface. This interface will ensure the turnover of the link from the satellite channel to the terrestrial channel at the command of the handover procedure protocol. The handover execution assistance required in the chosen inter-segment handover scheme, is provided by this entity. The test-bed will be designed for user bit rates up to 64 kbit/s, and the multimedia application proposed for the trials will be connected to the SINUS test-bed equipment through an ISDN interface.

4 Joint trials possibilities

In line with the spirit of European projects such as ACTS, one of the goals in the SINUS project is to contribute with research results and test equipment which may be of profit to other projects within the program. In a similar manner, the SINUS consortium would be most interested in extending the demonstrations by profiting from the experience and equipment development of other projects. Some ideas on joint trial possibilities with other ACTS projects have been presented both on short term within the actual lifetime of the project, and also as possibilities for future co-operations in an eventual extension of the concerned projects. Information has been exchanged with the FRAMES and the RAINBOW projects in order to harmonise the radio access mode and the implementation of the GRAN principle respectively. Joint trials are planned with the MOMUSYS (cfr. article in this magazine issue) and TOMAS projects. MOMUSYS can provide a prototype multimedia terminal for the SINUS trials, and the MOMUSYS consortium is also interested in the possibilities of multimedia transmission over the satellite channel. The TOMAS project will have access to an operating geostationary satellite, which will permit SINUS to test our connection over at least one real satellite constellation, and TOMAS is interested in trying out their terminal on our mobile satellite channel emulator. In an extension phase of the two projects, joint trials with RAINBOW seem attractive. The RAINBOW project is the equivalent to the SINUS project but for terrestrial UMTS, which means that by performing tests with both S- and T-UMTS demonstrators we will cover most of the integrated system configurations. The realisation of these joint trials may prove to be a complex task, this is why they are not considered possible within either projects' actual framework.

5 Conclusions and perspectives

The work within the SINUS project is about to conclude the specifications constituting the first phase of the project. The second phase will mainly be concerned with the development and integration of the test-bed entities. Finally, the last months of 1998 will be dedicated to the SINUS demonstrations in Torino, Italy.

The SINUS project has been particularly concerned with the specificities of a satellite component for communication within the UMTS concept, but has also tried to achieve architectural commonalities with T-UMTS through the adoption of the GRAN principle. The QoS requirements of S-UMTS have been specified in the first phase of the project, and the development following in the second phase is intended to meet those requirements. Particular attention has been given to the aspects concerning functionalities such as handover and connection establishment. Tests defined in the SINUS project will be performed, but co-operation with other projects permit interesting demonstrations, such as for instance testing of a newly developed multimedia terminal from the MOMUSYS project. Many other perspectives are open for future trials, both in the direction of real environment tests when mobile satellite systems are made available, or in the direction of enhanced test-bed demonstrations including other parts of a global system such as for instance emulation of the terrestrial channel. The main challenge of this project is however to show the possibilities and limitations of an integrated S-UMTS component in order to contribute to the standardisation and development of a European concept which may again influence the development on a global scale.

References

1 ITU-8.1/157-E. *Modularity for IMT-2000/FPLMTS*, RCSG, 10 February 1997.

2 ETSI- 23.05. *UMTS network principle,* SMG.

List of abbreviations

ACTS	Advanced Communications Technologies & Services
ATM	Asynchronous Transfer Mode
B-ISDN	Broadband Integrated Services Digital Network
CDMA	Code Division Multiple Access
FES	Fixed Earth Station
FRAMES	Future Radio Wideband Multiple Access Systems
GEO	Geo-stationary Earth Orbit
GRAN	Generic Radio Access Network
HEO	High Elliptic Orbit
IN	Intelligent Network
ITU	International Telecommunication Union
LAC	Link Access Control
LEO	Low Earth Orbit
MAC	Medium Access Control
MEO	Medium Earth Orbit
MOMUSYS	Mobile Multimedia Systems
MT	Mobile Terminal
N-ISDN	Narrowband Integrated Services Digital Network
QoS	Quality of Service
RAINBOW	Radio Access Independent Broadband On Wireless
SINUS	Satellite Integration into Networks for UMTS Services
S-UMTS	Satellite UMTS
TOMAS	Inter-trial Testbed of Mobile Applications for Satellite Communications
T-UMTS	Terrestrial UMTS
UMTS	Universal Mobile Telecommunication Systems

Vendela Paxal is Research Scientist in the Satellite & Radio Systems Unit at Telenor Research & Development, Kjeller

email address:
vendela.paxal@fou.telenor.no