Human Factors guidelines for multimodal interaction, communication and navigation

Matthias Schneider-Hufschmidt and Lutz Groh, Siemens;
Patrick Perrin, Actimage; Nick Hine, Uni. Dundee;
Stephen Furner, BT Exact Research.

Abstract

This paper identifies key issues, solutions and actions for multimodal interaction, communication and navigation at the user interface with ICT systems and terminals. It specifically addresses the usage context of transactional interactions for independent living, and focus on the dynamics of multimodal transactions/user dialogues for the consumers of ICT systems and terminals. It identifies how simplifications, translations, sensory transpositions, or other presentation or content manipulations of a multimodal transaction can be used to improve ease of access to telecommunications products and services for people with sensory, motor or cognitive impairments.

Consultation with users and user groups were carried out to identify the areas of transactional interactions that are currently providing a barrier to ICT access. To overcome these barriers a structured set of design and implementation principles is presented. The principles given here have been produced from a "design for all" perspective. While they have been produced out of observation and consultation with disabled people they have been constructed to be of value in making multimodal dialogues easier and more effective for all users not just those with special needs.

Key words: ICT, disability, multimodal, design for all, interaction, interface, diversity, user, ETSI, standardisation, guidelines

1. Multimodality for reduction of social exclusion

1.1 The social exclusion of disabled people

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

Being disabled, so it seems, still goes hand in hand with poverty caused by unemployment. Reliable statistics on unemployment among disabled people are hard to collect due to the fact that European societies employ different definitions of disabilities and use different methods for calculating the unemployment rate. Some evidence, however, suggests that employment rates among disabled people are low, at around 40 % [1] (compared to 65 % of the non-disabled

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

population), and half of those disabled people who have employment earn incomes below half the general population mean (often taken as an indicator of poverty)[2].

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

1.2 Monomodality of information presentation and social exclusion

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

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

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

Not being able to use a device or service because its input and output channels support one modality only is a serious restriction of one's everyday life. For example, it can mean that online banking becomes impossible and transactions have to be made personally (which may present other additional barriers). Information provided in trains or buses by voice or visual indication only seriously restricts those who cannot see or hear.

Similar restrictions exist for large sections of educational programmes. These cannot be used by disabled people thus limiting their options for professional promotion. Many jobs are designed around a technology that is not designed to provide multimodal interfaces - a large portion of working population is working with technology that requires seeing and hearing. The impact of monomodality on a person's social life is also dramatic. A hearing impaired person watching a movie is excluded from the sound, a visually impaired person from the picture (often friends or members of the family compensate by verbally explaining to a blind person or by signing to a deaf person).

1.3 Multimodality as compensation for impairments

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

---

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

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

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

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

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

1.4 Alleviating social exclusion

The need for making an increasingly knowledge-based society accessible to everybody, including disabled people, is acknowledged by national and international governmental and regulatory institutions[4]. The insufficient presence of multimodality in our society is one of the factors contributing to the social exclusion of disabled people. Multimodality can close the gap for many disabled people, making the world and its opportunities more accessible to them and thus contribute to a reduction of the social exclusion of disabled people in their social and professional life. Introducing multimodality is a necessary, but by no means sufficient first step for overcoming social exclusion.

2. Human factors guidelines and recommendations for multimodal interaction

The guidelines given here describe various aspects of multimodal interfaces. These guidelines were developed by the European Telecommunications Institute (ETSI) Special Task Force (STF) 204 with the support of the European Commission eEurope initative. Full details of the results of this activity can be found in ETSI Guide EG 202 191 V1.1.1 (2003-08) Human Factors (HF): “Multimodal interaction, communication and navigation guidelines”. They may be used to get answers to the question such as:

- In which application areas can the rules be applied, e.g. e-government, e-work, e-education, e-commerce, teleshopping, e-leisure, e-banking, e-care, e-health)?

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

- What are the representational building blocks used to create multimodal user interfaces?
- What are the transactional building blocks used to create complete transactions, e.g. authorization, data-entry, confirmation, searching, data presentation?
- How do abilities of different users affect the selection of modalities, where and when is the selection done?
- What are the effects of aesthetics, social acceptability, life style, pleasurenomics, branding and product differentiation on multi-modal user-interfaces?

2.1 Metaprinciples for the design of multimodal systems

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

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

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

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

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

2.2 Design principles for multimodal systems

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

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

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

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

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

Principle 3: ***Individual modalities should be activated optionally.***

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

Principle 4: ***Individual modalities should be scalable by the user.***

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

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

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

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

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

Principle 7: ***The user-specific modality setting should persist.***

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

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

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

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

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

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

Description: The interaction style as well as the style of content representation in multimodal systems should be consistent across different modalities. Transactional systems are being built for users to achieve certain goals and should support these activities through interfaces making the experience fun and the system a pleasure to use. Designers know about the importance of aesthetics and fun for the success of a system. These factors should be carried across to other modalities if a user is unable to use a specific modality for the interaction with the system.

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

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

Principle 12: Multimodality should not be stigmatizing.

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

2.3 Implementation principles

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

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

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

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

3. Conclusions and recommendations for future work

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

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

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

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

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

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

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

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

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

4. Acknowledgements

These guidelines were developed by the European Telecommunications Institute (ETSI) Special Task Force (STF) 204 with the support of the European Commission eEurope initative. Full details of the results of this activity can be found in ETSI Guide EG 202 191 V1.1.1 (2003-08) Human Factors (HF); Multimodal interaction, communication and navigation guidelines. We would like to acknowledge the help and support of the ETSI secretariat in this work and to thank the disabled people and care staff who gave generously of their time for consultation activities.

5. References

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

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

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

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