

---

BY Željko Obrenović, Julio Abascal, AND  
Dušan Starčević

# UNIVERSAL ACCESSIBILITY AS A MULTIMODAL DESIGN ISSUE

*Creating adaptable solutions to address diverse users  
and situations.*

**I**n recent years, many research activities have focused on design that aims to produce universally accessible systems, taking into account special needs of various user groups. These special needs are associated with many user factors, such as impairments of speech, hearing or vision, cognitive limitations, aging, as well as with various environmental factors [3]. Fields that address this problem, such as Usability, Universal Accessibility, Universal Design, or Inclusive Design [8] have been developed as relatively independent domains, but they share many aspects with other human-computer interaction (HCI) disciplines. However, researchers and practitioners are often not aware of interconnections among concepts of universal accessibility and “ordinary” HCI. In view of this situation, in this article we show there is a fundamental connection between multimodal interface design and universal accessibility, and that awareness of these links can help both disciplines. Researchers from these areas may use different terminology, but the concepts they use often have essentially the same meaning. We propose a unified conceptual framework where these areas can be joined.

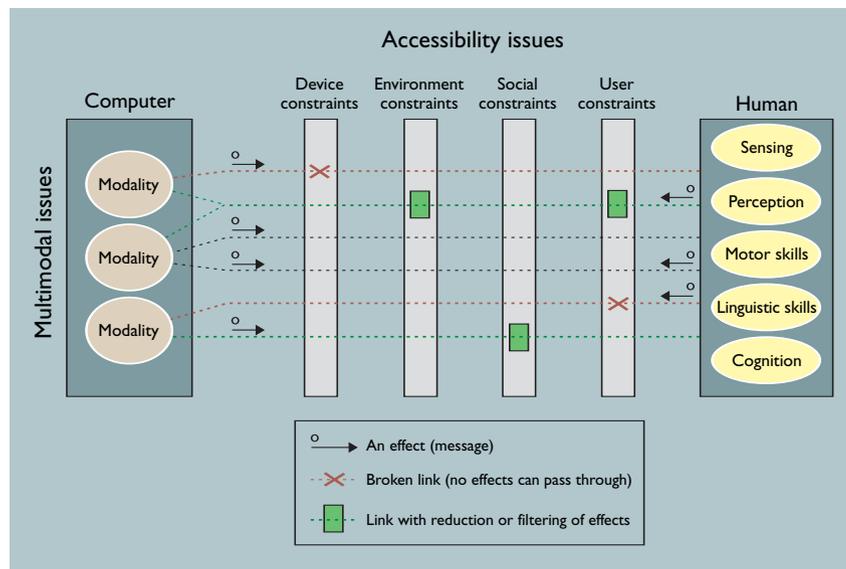
## ACCESSIBILITY AND MULTIMODAL INTERACTION

Universal accessibility and related approaches such as “Inclusive Design” or “Design for All,” aim to produce systems that can be used by everyone, regardless of their physical or cognitive skills. As this design philosophy tends to enhance the usability of the product, it can also be extremely valuable for non-disabled users trying to use the system under suboptimal conditions [3]. The activities resulting from the growing interest in accessibility and universal usability have produced a collection of resources that developers can use in their work. For example, many guidelines about accessibility, especially for Web design, are already available [10]. In addition, conferences such as the ACM Conference on Universal Usability (CUU), and the ACM Conference on Assistive Technologies (ASSETS), as well as journals such as the *International Journal on Universal Access in the Information Society*, offer good sources of practical and theoretical work in this area. Developers can also use various practical solutions and tools, such as Web site compliance checkers, semiautomatic Web site repair tools, or Web adaptation facilities that transform existing Web content “on the fly.” There are also activities in developing tools that use guidelines to automatically verify Web accessibility [1].

Multimodal interaction is a characteristic of everyday human discourse, in which we speak, shift eye gaze, gesture, and move in an effective flow of communication. Enriching HCI with these elements of natural human behavior is the primary task of multimodal user interfaces. Many studies have explored multimodal interaction from different viewpoints [2]. Sharon Oviatt gave a practical definition of multimodal systems, saying they combine natural human input modalities—such as speech, pen, touch, hand gestures, eye gaze, and head and body movements—in a coordinated manner with multimedia system output [6]. Matthew Turk and George Robertson further refined the difference between multimedia and multimodal systems, saying multimedia research focuses on the media, while multimodal research focuses on the human perceptual

channels [9]. Multimodal interfaces can improve accessibility for diverse users and usage contexts, advance performance stability, robustness, expressive power, and efficiency of communication [6].

While multimodal interaction research focuses on adding more natural human communication channels into HCI, accessibility research is looking for substitute ways of communication when some of these channels, due to various restrictions, are of limited bandwidth. What makes a difference between



**Figure 1. Modalities, constraints, and effects.** Computers and humans establish various communication channels over which they exchange messages with associated effects. Modalities process or produce these effects, while various interaction constraints reduce or completely eliminate some of these effects.

these two areas is a focus of their research. Therefore, many things from both areas can be generalized so that we can obtain a unified and more abstract view of them. In this way existing solutions from one domain could be applied in another.

## THE UNIFIED FRAMEWORK

Treating user interfaces as multimodal systems can clearly help design for universal accessibility, as multimodal interfaces describe HCI in terms of modalities, for example, in terms of communication channels established between the computer and the user. Environmental constraints or limited user abilities can be considered a break or decrease of throughput in these channels (see Figure 1).

If we describe user interfaces as a set of communication channels, and connect these descriptions with user, environment, and device profiles that describe limitations in usage of these channels, we can easily see if the multimodal interface will be appropriate for the user in a specific situation. However, to create a unified view of multimodal system design and accessibility, we need a semantic framework where we can explicitly and formally establish relations among concepts from both domains. Therefore, our first step is

to formally define a unified modeling framework for the description of multimodal HCI and various user and environment characteristics using the same terms. The proposed framework does not define any specific interaction modality—such as speech, gesture, graphics, and so on—nor a constraint, such as low vision, immobility, or various environment conditions, but defines a generic unified approach for describing such concepts. The framework, therefore, focuses on the notions of an abstract modality and abstract constraint, defining their common characteristics regardless of their specific manifestations.

**The Model of Multimodal HCI and Constraints.** Our approach is based on the idea that user interfaces can be viewed as one-shot, higher-order messages sent from designers to users [7]. While designing a user interface, the designer defines an interactive language that determines which effects and levels will be included in the interaction. Therefore, we model user interfaces with modalities they use, where we define a modality as a form of interaction designed to engage a number of human capabilities, for example, to produce effects on users, or to process effects produced by the user (see Figure 2). In our model, modalities can be simple or complex: a complex modality integrates other modalities to create simultaneous use of them, such as to provide modality fusion of fission mechanisms, while a simple modality represents a primitive form of interaction. Here, we do not focus on a detailed description of multimodal integration, but on the high-level effects a modality system or some of its parts use. We define input and output types of a simple modality, using the computer as a

reference point. An input modality requires user devices to transfer human output into a form suitable for computer processing. We classify these into event-based and streaming-based classes. Event-based input modalities produce discrete events in reaction to user actions, such as user input via a keyboard or mouse. Streaming-based modalities sample input signals with some resolution and frequency, producing a time-stamped array of sampled values. We introduce a special class of streaming modality, a recognition-based modality, which processes streaming data, searching for patterns. An output modality presents data to the user, and this presentation can be static or dynamic. A more elaborate description of this model can be found in [4].

While we describe HCI in terms of modalities, we describe various accessibility issues in terms of interaction constraints (see Figure 3). Interaction constraints can be viewed as

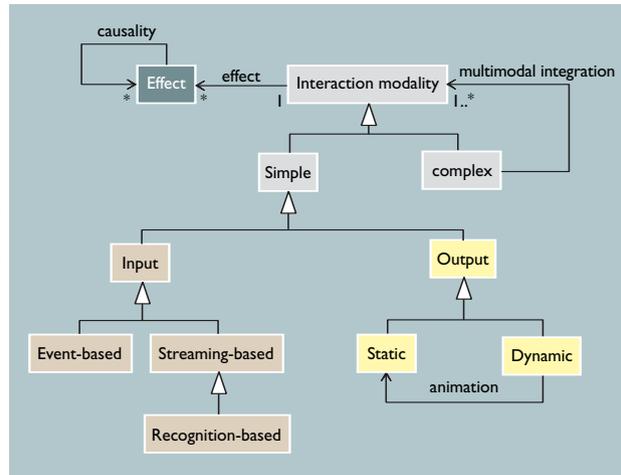


Figure 2. Simplified model of computing modalities.

the filters on usage of some effects. Constraints are organized as basic and complex. We identify two types of basic constraints: user constraints and external constraints. User constraints are classified into user features, states, and preferences. User features describe the long-term ability of a user to exploit some of the effects, and this description can include some of the user disabilities, such as low vision or immobility. A user state constraint, further classified in emotional and cognitive context, describes a user's temporary ability to use some effects. User preferences describe how much the user is eager to make use of some effects—it is a user's subjective mark of the effects they prefer or dislike.

External constraints are categorized as device constraints, environment constraints, and social context. Device constraints describe restrictions on the usage

The proposed model allows the flexible definition of various simple and complex constraints of different types. THE RESULTING CONSTRAINT IN A PARTICULAR SITUATION WILL BE A COMBINATION OF THE USER'S STATE, ABILITIES, AND PREFERENCES, as well as various external factors relevant to that situation.

of effects that are a consequence of device characteristics. For example, a mouse is limited to capture movement in two-dimensional space with some resolution, while output devices, such as screens on PDAs and other mobile devices, have limited resolution and a

ability and Health (ICF). Our model also allows developers to use other categories. In our model, these concepts are subclasses of the Effect class presented in Figures 2 and 3. Sensory effects describe processing of stimuli performed by human sensory apparatus. Per-

ceptual effects are more complex effects that the human perceptual system obtains by analyzing data received from sensors, such as shape recognition, grouping, highlighting, or 3D cues. Motor effects describe human mechanical action, such as hand movement or pressure. Linguistics effects are associated with human speech, listening, reading, and writing. Cognitive effects take place at a higher level of human information processing, such as memory processes or attention.

Effects are often interconnected. For example, all perceptual effects are a consequence of sensory effects. These relations among effects are important because in this way designers can see which side effects would be caused by their decisions to use particular effects.

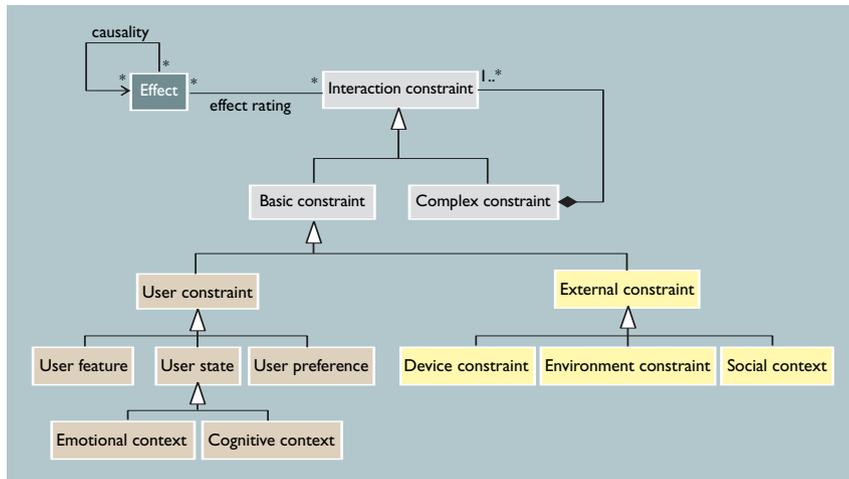


Figure 3. Simplified model of interaction constraints.

limited number of colors. Environmental constraints describe how the interaction environment influences the effects. For example, when driving a car, in most situations, users are not able to watch the screen and, therefore, this situation greatly reduces the usage of visual effects. In addition, various other environmental factors, such as visual conditions or noise, greatly affect the usage of other effects. Social context describes the social situation in which the interaction occurs. The proposed model allows the flexible definition of various simple and complex constraints of different types. The resulting constraint in a particular situation will be a combination of the user's state, abilities, and preferences, as well as various external factors relevant to that situation.

**Common Ground: The Effects.** Entities that connect modalities and constraints in our model are effects. We have classified effects used by modalities and affected by constraints in five main categories [4]: sensory, perceptual, motor, linguistic, and cognitive effects.

These categories are based on various sources, such as the World Health Organization International Classification of Functioning, Dis-

Table 1. Some example modalities and their associated effects.

#### USING THE FRAMEWORK

Our unified framework can be used to describe various interaction modalities and interaction con-

Modality	Composed of	Effect which the modality uses	Effect type	
Simple textual presentation	Pixel	Visual sensory processing	sensory	
		Letter		Correct central field vision
		Normal vision sharpness		
	Word	Shape recognition of grouped pixels	perceptual	
		Grouping of letters by proximity		
		Shape recognition of words		
Text line	Grouping of words by good continuation	-		
Paragraph	Grouping of lines by proximity Highlighting by the shape of the first line			
Aimed hand movement	Hand movement (input modality)	Hand movement	motor	
		Pressure	perceptual	
	Visual feedback (cursor)	Highlighting by shape of the cursor		
		Highlighting by motion		
		Highlighting by depth (cursor shadow)		
Visual menu	Selection	See aimed hand movement.	-	
	Menu presentation	Grouping by surrounding of menu borders	perceptual	
		Grouping of items by proximity		
		Highlighting by shape and color (selected item)		
		Visual reading of item text	linguistic	
		Understanding of the menu language		
Simple speech interaction	Speech input	Speaking	linguistic	
	Speech output	Listening	cognitive	
		Attention		

straints. By combining these descriptions, and by using effects as a common ground, it is possible to see if the designed interface will be appropriate for a particular situation, and it can enable adaptation of user interfaces according to user profiles and situational parameters.

**Describing User Interfaces, Users, and Environment.** We model user interfaces with their modalities, where we describe each modality with the effects it requires in order to be operative. For example, Table 1 shows effects produced by some common modalities, such as simple text presentation, aimed hand movement, visual menu interaction, and speech-based user interfaces, where we also illustrate different effect levels.

A specific user interface can then be described using these high-level descriptions of modalities, while we can obtain detailed descriptions of used effects automatically through mappings, such as those shown in Table 1. It is also possible to have several alternative mappings among modalities and effects according to different theories. For example, simple

textual presentation in Table 1 is described according to Gestalt psychology, but it is also possible to provide a description of these modalities according to other theories.

Disability	Effects reduced by a disability
Blindness	Absence of all visual stimulus processing
Poor acuity (poor sharpness)	Reduced visual sharpness
Clouded vision	
Tunnel vision	Reduced peripheral vision
Central field loss	Reduced central vision
Color blindness	Reduced color sensation and contrast processing
Deafness	Absence of all audio stimulus processing
Hard of hearing	Reduced audio sensory processing
Weakness	Reduced movement and pressure
Limitations of muscular control	
Limitations of sensation	Reduced pressure
Joint problems	Reduced movement
Pain associated with movement	
Dyslexia	Reduced linguistic effects
Attention Deficit Disorder	Reduced attention
Memory Impairments	Reduced memory processes

**Table 2. Some example disabilities and the associated constraints of effects.**

effects that are different from a typical user. Alternatively, it is possible to define a repository of user ability categories, where each category is described with a set of effects that it reduces. These categories can describe some disability, or other factors, such as average abilities of different age groups. For example,

**Table 3. Driving a car as a complex interaction constraint, composed of various simpler constraints, with associated effects.**

Constraints	Situations (constraint subclasses)	Influence on the usage of effect
Traffic situation (environmental constraint)	Car stopped	No specific reductions.
	Normal traffic situation	It is not convenient to require the user to <i>move</i> or <i>use hands</i> . Also, user's <i>central field vision</i> is directed toward the road.
	Traffic jam	In addition to the normal traffic situation, additional limitation is usage of user <i>attention</i> , as the user is more focused and stressed.
Noise level (environmental constraint)	Insignificant noise level	No specific reductions.
	Normal noise level	A user's <i>audio perception</i> , <i>audio 3D cues</i> , and <i>speech</i> can be used provided that they are of significant intensity.
	High noise level	<i>All audio effects</i> are significantly reduced.
Visual conditions (environmental constraint)	Day	No specific reductions.
	Night	Driving conditions are tougher; user is more focused and stressed.
	Fog	
Weather condition (environmental constraint)	Dry	No specific reductions.
	Rain	Driving conditions are tougher; user is more focused and stressed.
	Snow	
User current state (emotional context)	The driver is relaxed	No specific reductions.
	The driver is stressed	Limited usage of attention requests and complex interaction modalities.
Number of people in the car (social context)	The driver is alone	No specific reductions.
	The driver is not alone	<i>Other user</i> can use the application. Can affect the noise level and linguistic effects.

Accessibility issues, such as user abilities and environmental conditions, are described using interaction constraints. User abilities can be described in several ways. For example, one approach is to create individual profiles of each user, associating all the effects with values describing the user's capability to exploit them. For simplicity, the profile could include only the

effects that are reduced by some disabilities. In modeling and analyzing user interfaces, a very important role has the relations among effects. For example, if we describe that the user is not capable of processing a sound, it means not only sensory, but also all the audio perceptual effects will not be appropriate for that user.

In a similar way we can describe constraints introduced by the environment conditions. For example, driving a car is a complex constraint that integrates different user and environmental parameters. Table 3 shows a simplified description of this constraint that depends on traffic circumstances, weather conditions, noise level, lighting, user current state, as well as the number of people in the car. Constraints can also be interconnected, for example, visual and weather conditions can affect the user's current state, while the number of people in the car can influence the noise level. This

description can be useful in determining which modalities to use in particular situations. When the car is stopped, it is possible to use the central field vision of the user, as well as other effects to a greater degree. On the other hand, traffic congestion further limits possible usage of these effects, allowing their use to a lesser degree.

**Analysis and Transformations.** The descriptions of multimodal interfaces and interaction constraints presented here can be used for various purposes. In a simpler form, they can serve as metadata about some user interface, or as a part of a user profile. However, with formal descriptions of user interfaces and constraints, it is possible to develop tools that analyze and transform the content in order to see if it is suitable for particular situations or users.

We are developing a set of design-support and educational tools based on the proposed framework. These tools take as an input the description of a user interface, expressed in terms of modalities, and then evaluate it, for example, giving the list of effects, or giving a list of potential problems in some environments, as well as the list of user groups that could have difficulties using this interface. To increase the awareness about importance of accessibility aspects, these reports also contain data about the percentage of people who suffer from some interaction limitations (for example, approximately 8% of men and 0.4% of women have some form of color blindness).

Various other applications, such as dynamic adaptation and content repurposing are also possible. By connecting descriptions of user interfaces, user profiles, and other constraints, we can analyze and transform content in various ways. The proposed framework can, therefore, be a good basis for adaptation and content repurposing that attack the problem of developing content for various users and devices. A more detailed description of our previous work in this area can be found in [5].

## CONCLUSION

Our proposed approach can provide developers and researchers several advantages. From the developer's point of view, one advantage is that it is possible to design more flexible and more reusable solutions, aimed for a broader set of situations. Most of the previous work in designing solutions for people with disabilities concentrated on a specific set of disabilities, or on specific situations. Considering the great diversity of disabilities and situations, it is clear that development and maintenance of such systems is rather complex. With our approach, developers can concentrate on more generic effects, providing solutions for different levels of availability of specific

effects. In this way it is possible to create adaptable solutions that adjust to user features, states, preferences, and environmental characteristics.

Another important advantage is that our framework enables treating different situations in the same way. As user features and preferences are described in the same way as environmental characteristics, it is possible to use solutions intended for a user with some disability for a non-disabled user in situations that limit the interaction in the same way as some disability limits the other user. In addition to providing more universal solutions, this could also solve some of the associated problems because design is not concerned with disabilities (recognizing that the term "disability" often creates negative reactions), but with various effects and their constraints. **■**

## REFERENCES

1. Abascal, J. et al. The use of guidelines to automatically verify Web accessibility. *Universal Access in the Information Society* 3, 1 (2004), 71–79.
2. Blattner, M.M. and Gliner, E.P. Multimodal integration. *IEEE Multimedia* (Winter 1996), 14–24.
3. Nicolle, C. and Abascal, J., Eds. *Inclusive Design Guidelines for HCI*. Taylor and Francis, London, 2001.
4. Obrenovic, Z. and Starcevic, D. Modeling multimodal human-computer interaction. *IEEE Computer* 37, 9 (Sept. 2004), 62–69.
5. Obrenovic, Z., Starcevic, D., and Selic, B. A model driven approach to content repurposing. *IEEE Multimedia* 11, 1 (Jan. 2004), 62–71.
6. Oviatt, S., Darrell, T., and Flickner, M. Multimodal interfaces that flex, adapt, and persist. *Commun. ACM* 47, 1 (Jan. 2004), 30–33.
7. Prates, R., De Souza, C., and Barbosa, C. A method for evaluating the communicability of user interfaces. *Interactions* (Jan.–Feb. 2000), 31–38.
8. Savidis, A. and Stephanidis, C. Unified user interface design: Designing universally accessible interactions. *International Journal of Interacting with Computers* 16, 2 (2004), 243–270.
9. Turk, M. and Robertson, G. Perceptual user interfaces introduction. *Commun. of the ACM* 43, 3 (Mar. 2000), 33–35.
10. Web Accessibility Initiative (WAI); [www.w3c.org/wai/](http://www.w3c.org/wai/).

---

**ŽELJKO OBRENOVIĆ** (zeljko.obrenovic@cwi.nl) is a researcher at the Centre for Mathematics and Computer Sciences (CWI) in Amsterdam, The Netherlands. Material presented in this article is derived from his work in the Center for Command Information Systems in Belgrade, Serbia.

**JULIO ABASCAL** (julio.abascal@si.ehu.es) is a professor at the Department of Computer Architecture and Technology of the University of the Basque Country, San Sebastian, Spain.

**DUŠAN STARČEVIĆ** (starcev@fon.bg.ac.yu) is a professor in the School of Business Administration at the University of Belgrade in Serbia and Montenegro.

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.