

Chapter 5

Model-Based Evaluation: A New Way to Support Usability Evaluation of Multimodal Interactive Applications

Regina Bernhaupt¹, David Navarre², Philippe Palanque²
and Marco Winckler²

¹ ICT & S-Center, Universität Salzburg, Austria, e-mail: Regina.Bernhaupt@sbg.ac.at

² LIIHS-IRIT, University Paul Sabatier, France

Abstract Multimodal interfaces are becoming more common, even in the field of safety critical interactive software, mainly due to the naturalness of the interaction that increases the bandwidth between the user and the system they are interacting with. However, the specificities of multimodal interactive systems make it difficult to gather information from the use of modalities and to extract from this information recommendations for improving the multimodal user interfaces. This chapter aims at presenting how abstract information described in models can be fruitfully exploited to improve the quality of evaluations of multimodal interfaces. The approach presented in this chapter combines model-based verification (based on simulation scenario extraction generated from models) and empirical methods for usability evaluation. Our aim is to try to bring together two separated (and often opposite) issues, such as usability and reliability, into the development of safety critical systems. This approach is illustrated via a Space Ground System of a satellite control room, whose multimodal interaction technique is fully described by the means of formal models.

5.1 Introduction

Since the seminal work by Bolt (Bolt 1980) (Bolt & Herranz 1992), multimodal interaction techniques are considered a promising way to increase communication bandwidth between users and systems and to enhance user satisfaction and comfort by providing a more natural way of interacting with computer systems. Several studies have shown that using two pointing devices in a normal graphical user interface is a more efficient and understandable interaction than using basic mouse and keyboard (Buxton & Meyers 1986; Kabbash, Buxton & Seller 1994; Zhai, Barton & Selker, 1997). In addition to subjective factors like comfort and satisfaction, increasing communication bandwidth between users and systems can have a significant impact on efficiency. For instance, the number of commands triggered by the users within a given amount of time and the error rate—typically the number of slips or mistakes made by the users (Reason 1990)—are influenced by the user interface and

01 its interaction techniques. Besides, the complementary nature of modalities can be
02 used to reinforce and clarify the communication between the users and the system
03 (Oviatt 1999).

04 Nevertheless, multimodal interaction is not a panacea. Studies of Dillon and
05 colleagues (1990) and by Kjeldskov and Stage (2004) *unsurprisingly* revealed that
06 when multimodal interfaces are poorly designed they are neither better understood
07 nor more efficient than any other user interface offering more standard interaction
08 techniques. To determine the contribution of modalities to the user interaction, many
09 empirical studies have been carried out in terms of

- 10 • Showing how usability and user acceptance is influenced by new devices and
11 novel interaction techniques (Bowman, Gabbard & Hix 2002; Hinckley, Pausch,
12 Proffitt & Kassel 1998; Nedel, Freitas, Jacob & Pimenta 2003; Poupayev,
13 Weghorst, Billingham & Ichikawa 1998)
- 14 • Showing that the perceived usability is impacted according to the kind of tasks
15 performed (Dybkjær, Bernsen & Minker 2004; Jöst, Haubler, Merdes & Malaka
16 2005) and according to the context of use (e.g., indoor x outdoor conditions,
17 mobile applications) (Baille & Schatz 2005)
- 18 • Trying to assess the accuracy of multimodal interaction for given tasks (Balbo,
19 Coutaz & Salber 2003; (Kaster, Pfeiffer & Bauckhage, 2003; Suhm, Myers &
20 Waibel 1999; Holzapfel, Nickler & Stiefelhagen 2004)

21
22 Some of these investigations show that low-level captured data (e.g., users' events
23 such as mouse clicks and speech) and high-level users' intentions (e.g., goals and
24 tasks) must be combined to determine the accuracy and the perceived usability. This
25 is noteworthy that many users prefer multimodal interaction and nonconventional
26 input devices despite a poorer performance (Kaster, Pfeiffer & Bauckhage 2003).
27 The multimodal dimension brings additional interesting issues to usability evalua-
28 tion methods. Indeed, each element involved in the design of the user interface can
29 have a huge impact on its usability. For instance, results of existing empirical studies
30 of multimodal applications revealed intricate problems concerning the assessment
31 of the usability of a multimodal interface with respect to several dimensions such
32 as usage and interpretation of modalities, individual user preferences for modality,
33 context-of-use, choice of input and output devices, and interaction techniques.

34 Despite the fact that all these issues increase the difficulty of evaluating mul-
35 timodal interfaces usability, these interfaces are becoming more common even in
36 the field of safety critical interactive software such as military (Bastide, Navarre,
37 Palanque, Schyn & Dragicevic 2004) and medical systems (Trevisan, Vanderdonck,
38 Macq & Raftopoulos 2003). A safety-critical system is a system for which the cost
39 of a failure is significantly more important than the development costs. User error or
40 usability problems might have dramatic consequences, leading to loss of lives. This
41 kind of system requires thorough evaluation and testing to ensure both usability
42 and reliability. This chapter proposes a new approach that combines model-based
43 specification (typically used for the design of this kind of system) and empirical-
44 oriented methods for usability evaluation. The model-based approach relies on
45 formal description techniques and is used to support the assessment and usability
46 evaluation of the multimodal user interface. This combined approach addresses two

01 main drawbacks of current practice in the field of usability evaluation of multimodal
02 systems:

- 03 • Lack of support for a complete understanding of the detailed behavior of the
04 system (both at the level of interaction and at the dialog level). This problem can
05 be overcome by appropriate modeling support, because it is illustrated hereafter
06 via *A Case Study for a Space Ground System in a Satellite Control Room*
- 07 • Poor integration of usability results into the whole development process. This
08 issue is discussed in the subsection *Modifying Models to Accommodate Change*
09

10 The next section briefly presents the state of the art in the field of usability evalu-
11 ation of multimodal interfaces. We then informally present a case study for a Space
12 Ground System used in satellite control room that is fully described by means of
13 the interactive cooperative objects (ICO) formal description technique (Dragicevic,
14 Navarre, Palanque, Schyn & Bastide 2004). This case study is used in the rest
15 of the paper as a concrete example of multimodal interaction techniques applied
16 to safety critical systems. We briefly describe the results of usability evaluation
17 for this application with two traditional methods (i.e., usability test and cognitive
18 walkthrough). We then present the shortcomings of these two traditional methods,
19 and show how model-based evaluation could support these methods and reduce
20 the identified limitations. This support is mainly provided through the generation
21 (from the formal models) of usability evaluation scenarios that are then used in
22 standard usability evaluation methods. The last section details the advantages and
23 lessons learned from model-based usability evaluation. We show how model-based
24 usability evaluation extends current usability evaluation practice (especially as far
25 as multimodal interfaces are concerned), as well as the limitations of this approach.
26
27

28 **5.2 Usability Evaluation of Multimodal Systems**

29
30

31 This section briefly presents the state-of-the-art in the field of usability evaluation
32 of multimodal interfaces. We first present the peculiarities of such interfaces and
33 then compare the approaches that have been designed for supporting their usability
34 evaluation.
35
36

37 **5.2.1 Specificities of Multimodal Interactive Systems**

38
39

40 Even though multimodality is usually associated with the possibility for the user to
41 use several input devices, multimodality concerns both input and output.

42 A specific aspect of multimodal user interfaces is that interaction techniques,
43 input/output devices, and sensory channels are closely related. Table 5.1 shows that
44 an interaction technique can involve one or more input devices or device combi-
45 nations. For example, for the ray-casting interaction technique, the synergistic use
46 of data glove and position trackers can replace a 3-D mouse as an input device.
On the other hand, a single device can be used in several interaction techniques.

Table 5.1 Examples of input interaction techniques and devices

Sensory channels	Interaction Techniques	Input Devices
Acoustic	Speech	Computer microphone, Cell phone, handheld
Haptic	Typing	Keyboard, touch screen
	Direct manipulation	Mouse, 3D Mouse, touch screen, Panthon
	Gesture interaction	Mouse, 3D Mouse, Data glove, position trackers
	Ray casting	3D Mouse, Data glove + position trackers
Visual	Gaze interaction	Video camera (eye tracking)
	Optical tracker	Video camera (optical markers)

For example, a touch screen can be used as an input device for several interaction techniques such as typing (on a soft keyboard), direct manipulation, and gesture interaction.

To assess the usability of a multimodal application, it is mandatory to evaluate not only the user interface per se, but also to take into account the couple *device-interaction technique*. In the same way that designing the multimodal user interface requires the selection of the appropriate couple (device, interaction technique), the evaluation has to address this issue even though in many multimodal user interfaces redundancy (if made available) allows users to interact with the application in various ways to trigger the same command or to enter the same data. In such cases, the couple (device, interaction technique) selected by the user might differ from one user to another but also with the same user between two successive tasks or usages of the application.

Another major issue of multimodal system concerns the fission and fusion of modalities. This issue is addressed by the classification proposed by Coutaz, et al. (1995), which includes one or more uses and interpretations (i.e., exclusive, concurrent, alternating, and synergistic) of both input and output modalities. As stated in Coutaz, et al. (1995) on a multimodal user interface, input and output modalities can be combined in four different ways (called the CARE properties): Complementarity, Assignment, Redundancy, and Equivalence. CARE properties can therefore be used to structure the usability evaluation of multimodal application, but because the users will be able to choose any interaction technique available, usability evaluation scenarios have to specifically address this issue.

Table 5.2 Examples of output interaction techniques and devices

Sensory Channels	Interaction Techniques	Output Devices
Acoustic	Voice synthesis	Voice synthesizer
Haptic	Force feedback	Panthon, Cybergrasp, Cyberforce
Visual	Image display	Computer screen, touch screen,
		head-mounted display, stereo glasses

AQ1



01 **5.2.2 Usability Evaluation Methods (UEM) Used**
 02 **for Multimodal Interfaces**
 03

04 As introduced before, it is a requirement for usability evaluation methods to take
 05 into account the specificities of multimodal interfaces. This section presents various
 06 UEMs that have been applied to and customized for multimodal user interfaces eval-
 07 uation. Figure 5.1 structures this information in four main categories and provides a
 08 summary of the most representative methods in each category.

09 Most of the usability studies for multimodal interfaces exploit some user testing
 10 where users' activity is observed and recorded, while users are performing pre-
 11 defined tasks. User testing is a preferred strategy for evaluation as it allows the
 12 investigation of how users adopt and interact with multimodal technology, providing
 13 valuable information about both usability and user experience.

14 Several types of user testing have been conducted, both in usability laboratories
 15 and in field studies, revealing user preferences for interaction modalities based on
 16 factors such as acceptance in different social contexts—noisy and mobile environ-
 17 ments (Jöst, Haubler, Merdes & Malaka 2005). In the following sections, we will
 18 use a case study to show how user testing with log-file analysis and think-aloud
 19 protocols can be customized to address the needs of MMI evaluation.

20 Evaluation based on inspection methods assumes that human factors experts
 21 rely on ergonomic knowledge provided by guideline recommendations, or on their-
 22 own experience, to identify usability problems while inspecting the user interface.
 23 Known methods belonging to this category include *cognitive walkthrough* (Lewis,
 24 Polson & Wharton 1990; Polson, Lewis, Rieman, & Wharton 1992), formative

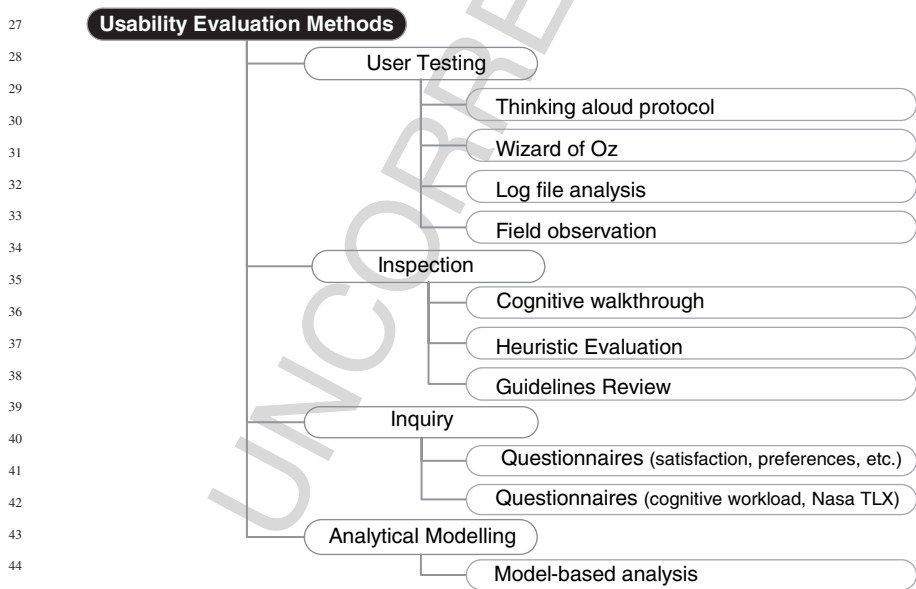


Fig. 5.1 An overview of evaluation methods

01 evaluation and heuristic evaluation (Nielsen & Mack 1994), and benchmarking
02 approaches covering issues such as ISO 9241 usability recommendations or confor-
03 mance to guidelines (Bach & Scapin 2003). Inspection methods can be applied in
04 the early phases of the development process through analysis of mock-ups and pro-
05 totypes. The lack of available ergonomic knowledge might explain why inspection
06 methods have been less frequently employed with an exception in Bowman, Gab-
07 bard & Hix (2002). Knowledge is not only missing in terms of experts' experience
08 for the design of multimodal systems, but also due to a lack of guidelines to cover all
09 potential modalities and modality combinations that might be encountered in mul-
10 timodal interfaces. Cognitive walkthroughs are designed to assess the achievement
11 of goals focusing on the goal structure of the interface rather than on interaction
12 techniques. We will show in the following sections how a cognitive walkthrough
13 can be used when evaluating a multimodal interface, and how this method must be
14 adapted to address the peculiarities of multimodal interfaces.

15 Questionnaires have been extensively employed to obtain qualitative feedback
16 from users (e.g., satisfaction, perceived utility of the system, and user prefer-
17 ences for modality) (Kaster, Pfeiffer & Bauckhage 2003) and cognitive workload
18 (especially using the NASA-TLX method) (Brewster, Wright & Edwards 1994;
19 Kjeldskov & Stage 2004; Trevisan, Nedel, Macq & Vanderdonckt 2006). Quite
20 often, questionnaires have been used in combination with user-testing techniques
21 as presented in (Jöst, Haubler, Merdes & Malaka 2005).

22 More recently, simulation and model-based checking of system specifications
23 have been used to predict usability problems such as unreachable states of the sys-
24 tems or conflict detection of events required for fusion. Paternò and Santos (2006)
25 propose combining task models based on concur task tree (CTT) notation with mul-
26 tiple data sources (e.g., eye-tracking data, video records) to better understand the
27 user interaction.

30 **5.3 A Case Study for a Space Ground System in a Satellite** 31 **Control Room**

32
33
34 This section presents a case study for a Space Ground System application to be used
35 in satellite control rooms (Ould, et al. 2004). The case study exploits multimodal
36 interaction techniques for the manipulation of a 3-D representation of a DEMETER
37 satellite, which stands for Detection of Electro-Magnetic Emissions Transmitted
38 from Earthquake Regions. More information about this satellite's functions and
39 missions can be found on <http://smc.cnes.fr/DEMETER/index.htm>.

42 **5.3.1 Informal Description of the Case Study**

43
44
45 This application provides multimodal interaction techniques to a user in charge of
46 moving the point of view (we later call this *navigating*) in a 3-D model of a satellite.

01 This navigation can be done either by rotating the 3-D model of the satellite directly,
02 using the mouse on the 3-D image, or using the two control panels presented in
03 Figure 5.2.

04 The control panel (b), entitled *point de vue* allows the user to manipulate the
05 current position of the point of view of the 3-D image using the set of buttons in the
06 top right hand side of Figure 5.2(b). The set of buttons in the Orientation subsection
07 allows one to rotate the satellite image in any direction. The set of buttons in the
08 Position subsection allows one to move the satellite image in any direction (up,
09 down, left, right, backward, and forward). The two list boxes on the left-hand side
10 present the list of components of the satellite and the list of categories the compo-
11 nents belong to, respectively. We do not present the other parts of the user interfaces
12 as they are related to functions beyond the scope of this paper.

13 In the initial state the satellite appears as presented in Figure 5.2(a). The main
14 task given to the user of this application is to locate one or several components
15 in the satellite. This task is not easy to perform because components are nested
16 and might not be visible (as they may be either behind or inside a component). To
17 support this task, the user interface makes it possible to set a transparency level for
18 the components' appearance from partly to fully transparent by selecting a percent-
19 age. This transparency is set by means of the Transparence slider on the right-hand
20 side of Figure 5.2(c). The goal of the user is to locate components that are either
21 overheating or overconsuming energy. The selection of the range of temperature of
22 interest can be done using the range slider in the *données* section on the right-hand
23 side of Figure 5.2(c). This part of the user interface can also be used for selecting
24 the energy consumption. Figure 5.3 presents a snapshot of the 3-D satellite model,
25 including the temperature of the visible components.

26 In this application, multimodal interaction takes place both while using the button
27 pairs (changing the point of view of the 3-D model), and while interacting with the
28 range slider (selecting the temperature and the consumption).

29 Due to space constraints, we only present multimodal interaction on the button
30 pair here. The interested reader can see the formal specification of a similar multi-
31 modal range slider component in Dragicevic, Navarre, Palanque, Schyn & Bastide
32 (2004), and the formal specification of a virtual chess game in Navarre, et al. (2005).

33 The controller's tasks are represented in Figure 5.4 using the CTT notation. We
34 only present here the tasks related to the management of functions that can be trig-
35 gered through multimodal interaction. The main goal of the controller is to monitor
36 the satellite. This goal is separated into three main tasks—monitoring temperature,
37 monitoring energy consumption, and locating physical components of the satellite
38 by moving its 3-D representation. It is important to see that the task model only
39 describes interaction at quite a high level because it is only describing what the
40 user is aiming at and not how to actually perform these tasks. Connection between
41 these high-level tasks and lower-level ones is done using the precise and complete
42 descriptions embedded in the system model. The next section shows precisely the
43 type of information embedded in the system model, as well as how this connection
44 is made.

45 Figure 5.5 shows an example of multimodal interaction for this application. On
46 this figure, the user is concurrently using three input devices—two mice and a

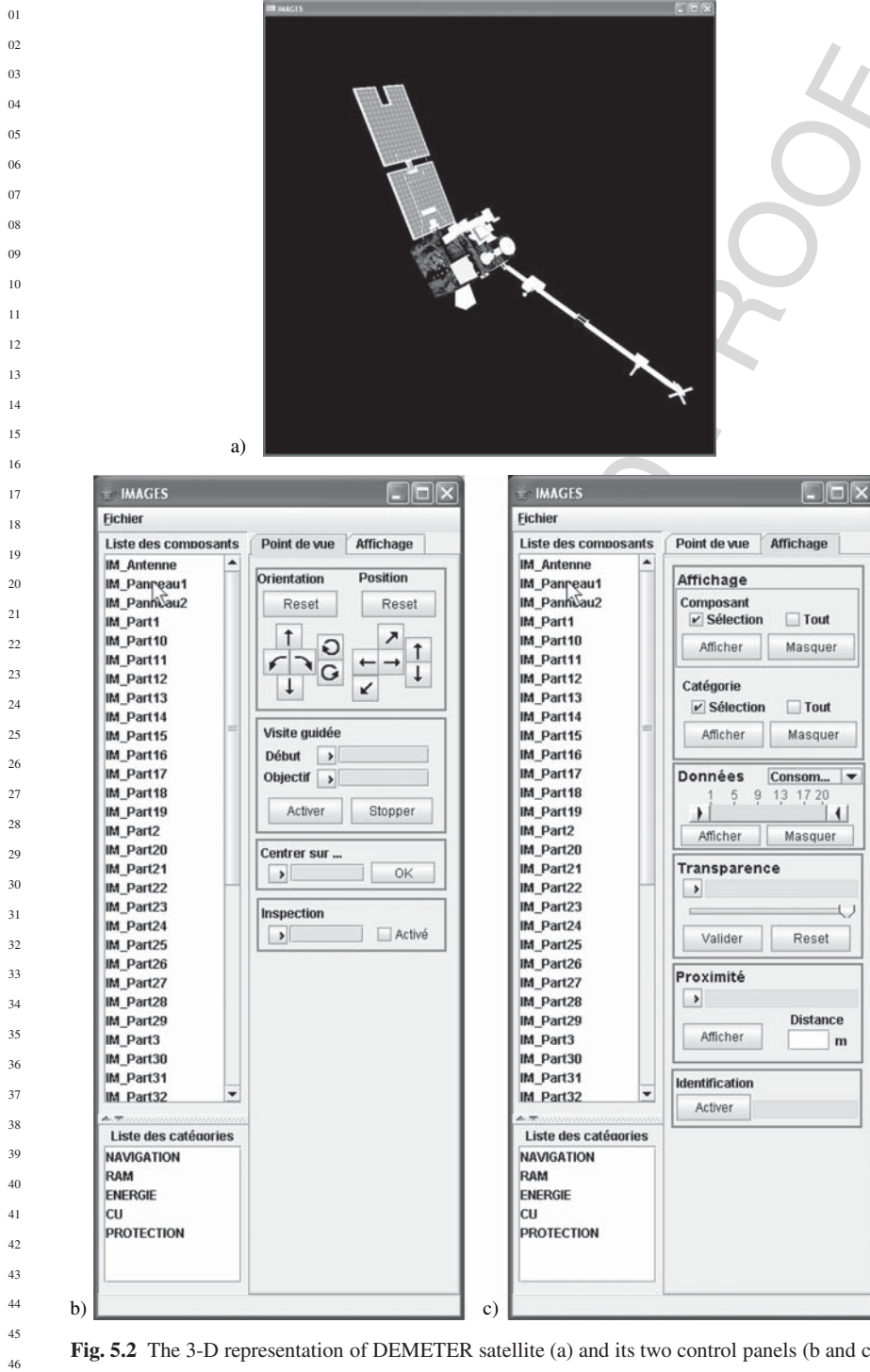


Fig. 5.2 The 3-D representation of DEMETER satellite (a) and its two control panels (b and c)

01
02
03
04
05
06
07
08
09
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

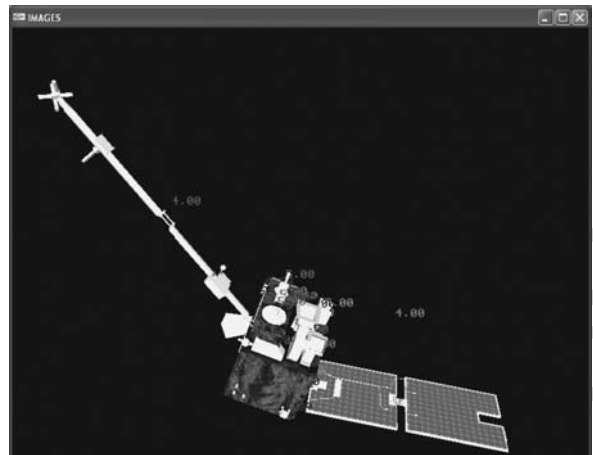


Fig. 5.3 3-D satellite model displaying the temperature of the visible components

speech recognition system. The speech recognition system only reacts to two different words: *fast* and *slow*. The interaction takes place in the following way—at any time, the user can use any of the mice to press on the buttons that change the point of view. In Figure 5.5, the button that moves the satellite image backwards (with the additional label right mouse interaction on Figure 5.5) has been pressed using the right mouse. Simultaneously, the left mouse is positioned on the button moving the satellite image to the left. At that time, the image has already started to move backwards, and as soon as the other button is pressed, the image moves both backwards and to the left. The user is also able to increase or decrease the movement speed by uttering the words *fast* and *slow*. In Figure 5.5, the word *fast* has been pronounced and recognized by the speech recognition system (as shown on the left-hand side of Figure 5.5). This action will reduce the time between two movements of the image. Indeed, the image is not moved according to the number of clicks on the buttons, but according to the time the buttons are kept pressed by the user.

Describing such interaction techniques in a complete and unambiguous way is one of the main issues to be solved, while specifying and developing multimodal interactive systems. The next section presents how the ICO formalism is able to deal with these issues. Additionally, it will show that the description above is incomplete

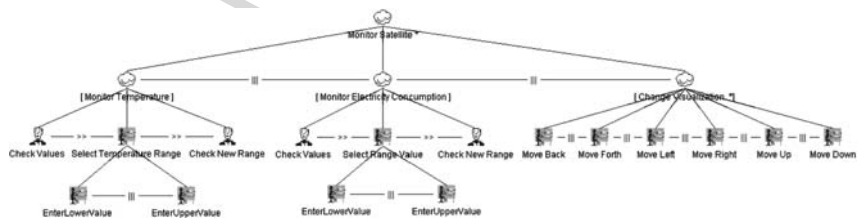


Fig. 5.4 CTT model of the tasks featuring multimodal interactions

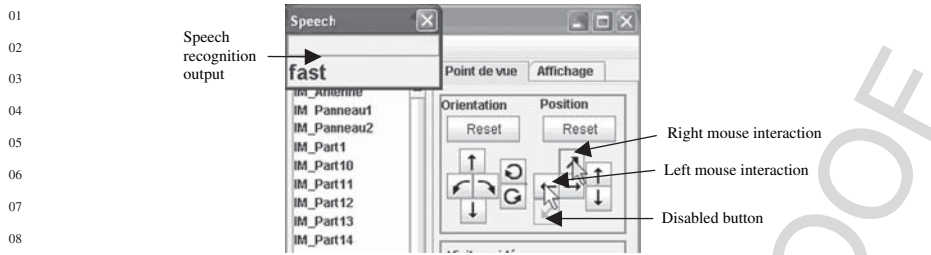


Fig. 5.5 One of the multimodal interactions in the application

and does not address (at an adequate level of detail) both timed and concurrent behavior, at least when it comes to implementation issues. This point is critical when usability evaluations are carried out. Indeed, to assess the comparative usability of two or more multimodal interaction techniques, a precise definition of test scenarios is required. This calls for tools and techniques that describe in a complete and unambiguous way interaction techniques at a very low level of detail. Such a technique is described in the next section.

5.4 Modeling of the Case Study

This section is devoted to the formal modeling of the multimodal interactive application presented in the previous section. In this multimodal application, there is no fusion engine, per se—the two mice are handled independently, and the speech interaction affects movement speed regardless of what interaction is performed with the mice.

The modeling is structured as represented in Figure 5.6. The right hand side of the figure shows the user interacting with the input devices. As stated before, three input devices are available. To configure this set of input devices, we use a dedicated notation called Icon (Dragicevic & Fekete 2001). A more readable model of this

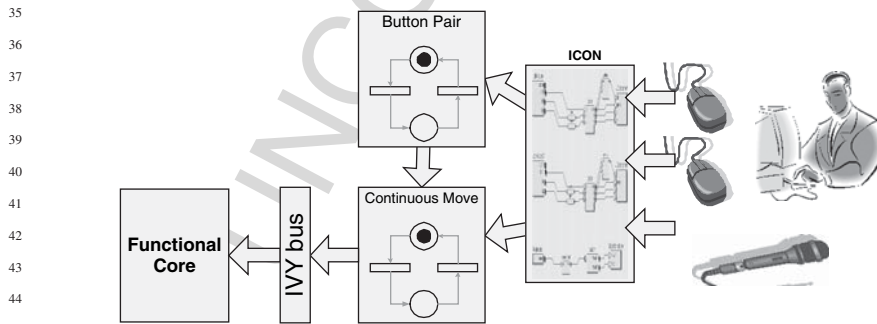


Fig. 5.6 Software architecture of the multimodal interactive application

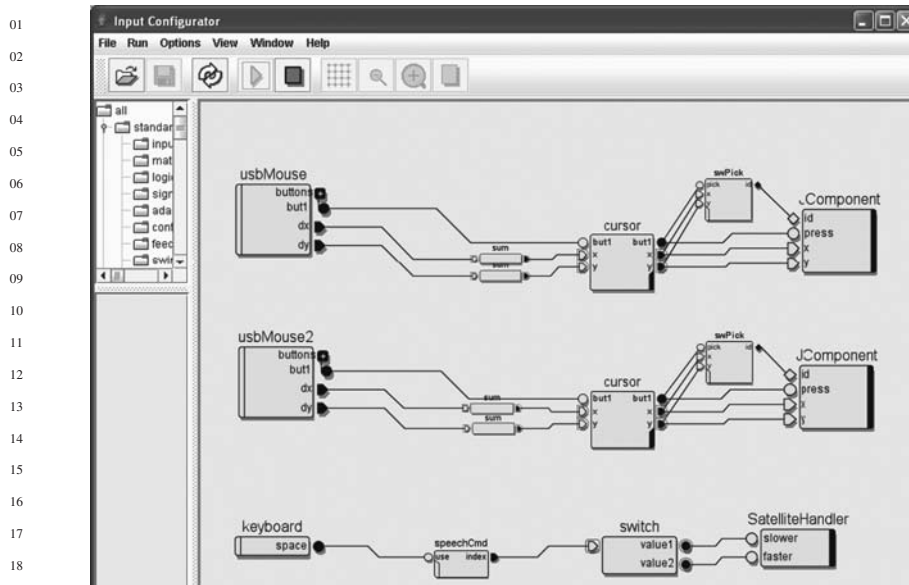


Fig. 5.7 Input configuration using Icon. (Dragicevic & Fekete 2001)

configuration is represented in Figure 5.7. This Icon model is then connected to two ICO models that are, in turn, connected via a communication bus called IVY to the functional core gathering all the data about the DEMETER satellite.

The left-hand side of Figure 5.7 represents the three input devices connected to software components. These components are represented as graphical bricks, and connectors model the data flow between these bricks. For instance, it defines that interaction with the mice will take place using the left button (but1 in the usbMouse brick), and that the alternate button for the speech recognition system is the space bar (Space label in the keyboard brick connected to the speechCmd brick). The right-hand side of this figure (Figure 5.7) represents contact points with the other models of the application. Because input configurations are not central to the scientific contribution of this paper, we do not present in more detail how this modeling works. More information about the system supporting the edition and execution of models, the behavior of a model, and the connections to other models can be found in (Navarre, Palanque, Dragicevic & Bastide 2006). Similarly, the functional core and communication protocol between the functional core and the interaction models are not presented.

The ICO model in Figure 5.8 represents the complete and unambiguous temporal behavior of the speech-based interaction technique, as well as how speech commands impact the temporal evolution of the graphical representation of the 3-D image of the satellite. Darker transitions are available according to the current marking of the models. Taking into account the current marking of the model of Figure 5.8 (one token in places delay, idle, and core), only three transitions startMove_1, faster_, and slower_ are available. These transitions describe the

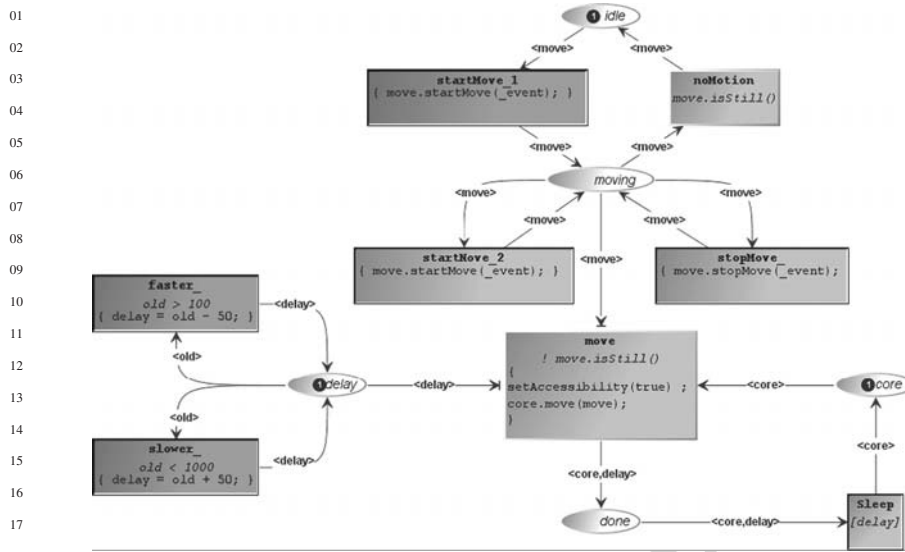


Fig. 5.8 Model of the temporal evolution of movements driven by speech (ContinuousMove class of Figure 5.6)

multimodal interaction technique available (i.e., how each input device can be used to trigger actions on the system). Transitions *faster_* and *slower_* are triggered when the user utters one of the two speech commands *fast* and *slow*. In the initial state, these are available and will remain available until the upper limit or the lower limit are reached ($delay > 1000$ for transition *slower_* and $delay < 100$ for transition *faster_*, respectively).

Figure 5.9 presents the model of the second ICO class of the application. This class is responsible for describing the behavior of each button pair. By button pair, we mean the buttons that are performing opposite actions like (up, down), (left, right) and (backward, forward). These three button pairs are represented on the right-hand side of Figure 5.5. To model these incompatible behaviors, the ICO description represents the fact that the user can press either the positive or negative button (e.g., *up* being the positive and *down* being the negative). Connection to the input device (the mice) is done using the Icon model of Figure 5.7.

5.5 Evaluation of the Case Study

Hereafter, we present the usability evaluations that have been carried out on the case study. Our aim with these evaluations is to describe how a formal description of multimodal interaction techniques can inform a traditional usability evaluation method (UEM)-like usability test and cognitive walkthrough. This approach, integrating both modeling techniques and usability evaluation, is hereafter called

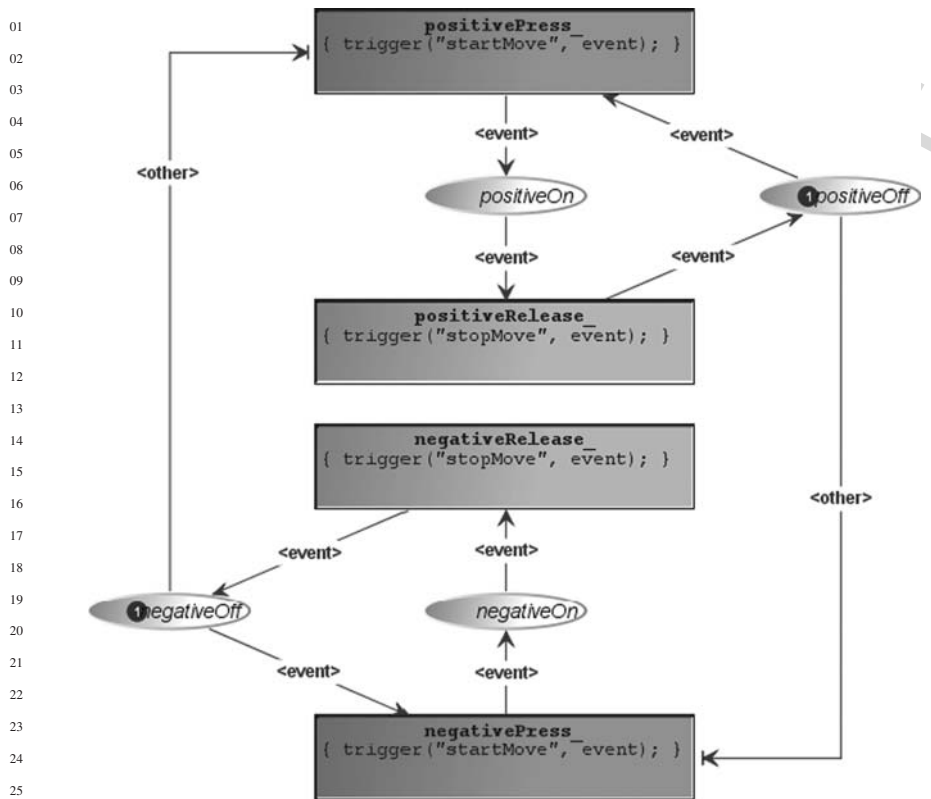


Fig. 5.9 Mutual exclusion of the pair of buttons for changing the point-of-view (ButtonPair Class in Figure 5.6)

model-based evaluation. Before explaining how such an integrated approach works, we present limitations of current UEM for addressing the specific issue of interaction techniques evaluation.

5.5.1 Limitations of Usability Test and Cognitive Walkthrough

The goal of a usability test is to identify major usability problems within the interface. While a common practice is to mainly use the most frequently performed tasks (extracted from task analysis, for instance), in the field of safety-critical systems, it is important to cover all (or most of) the possible interactions in which the user might be involved. The explicit description of the interaction techniques in the formal models makes it possible to identify not only the *minimum* number of scenarios to be tested, but also to select the tasks that are to be focused on more rationally. This selection can be done using analysis techniques on the models that will help

01 designers identify complex or cumbersome behaviors¹ that have been modelled and
02 might require more specific usability testing.

03 When testing multimodal interfaces, selection of scenarios reaches a higher level
04 of complexity due to the significant number of possible combinations of input
05 modalities, and also due to the fact that fusion engines usually involve quantitative
06 temporal evolution as shown in the ICO modeling of the case study. To test all (or
07 most) of these combinations, it is required to provide usability testing scenarios
08 at a much lower level of description than is usually done with systems featuring
09 more classic interaction techniques. Indeed, as for walk-up and use systems, the
10 interaction technique must be natural enough for the user to be able to discover
11 it while interacting with the system. In the field of safety-critical systems, training
12 and practice are essential points to be taken into account in the evaluation of
13 the system.

14 Even though we need to address this issue of low-level scenarios, it is also
15 important to notice that usability testing is very different from software testing
16 (which is usually dealing with those low-level tests). The objective here is to test
17 the usability of the interaction technique and not its robustness or default-freeness,
18 as in classical software testing. Software testing of the user interface is starting to
19 get attention from software engineering, but current solutions only deal with basic
20 WIMP interfaces (Memon, Pollack & Soffa 2000). The issue of reliability testing
21 of multimodal interactive systems is also very important, but is beyond the scope of
22 this chapter.

23 Formal description techniques can help to identify pertinent low-level interaction
24 scenarios and thereby inform selection of tasks more appropriately. Of course, the
25 number of scenarios is likely to be infinite (especially due to the number of possible
26 combinations of uses of input devices) but the model can support the identification
27 of equivalence classes of scenarios (i.e., the ones that are leading to the same state
28 changes in the system model).

29 To illustrate the advantages and drawbacks of model-based evaluation applied to
30 evaluate our Space Ground system, we present hereafter respective outcomes that
31 we can obtain from both standard usability evaluation (i.e., user testing in usability
32 labs and cognitive walkthroughs) and model-based evaluation.

33
34

35 **5.5.1.1 Standard Usability Evaluation**

36
37

38 User testing is typically performed in a laboratory, as shown in Figure 5.10 (some-
39 times in the field), where users are asked to perform selected tasks. The users are
40 observed by camera, and they might be asked to talk aloud (also called elicitation
41 activity) while performing the task. A usability test typically begins with the users

42
43

44 ¹ Our goal is not to go into detail about the definition of a cumbersome or complex models, but the
45 modeling constructs used within a model can provide such information. For instance, in Palanque,
46 Bastide & Paterno (1997), we have shown (in the field of Air Traffic Control) that multiple uni-
fications on incoming arcs of a transition (which is not a frequent phenomenon in models) might
result in tasks that are hard to perform.

01 **Fig. 5.10** Example of
02 usability test in action



03
04
05
06
07
08
09
10
11
12
13
14 answering a prequestionnaire related to the domain of the software (e.g., use of
15 other related systems, experience with multimodal-interfaces, hours of training, etc).
16 Some tasks are then performed to ensure that the user is able to use the system. The
17 experimenter (or test leader) usually describes the task to the user verbally and also
18 hands over a printed version of the task.

19 To test the task on identifying overheated components within the satellite, a
20 description might be as follows:

21 “Please find all components of the satellite with a temperature between 3° C and 6° C and
22 position the 3-D in such a way that the component with the highest temperature is visible.
23 Whenever you think that you would stop this task because you feel it is too complicated,
24 please tell us.”

25
26 The maximum time to solve this task is defined. The task is finished when the
27 user successfully solves the task, when the user takes more than the maximum time
28 to solve the task, or when the users states that he would give up and requires some
29 help. Several rounds of usability testing are performed with different users. The
30 number of successful completions and the completion time are recorded. Tasks not
31 solved indicate usability problems, leading to further detailed investigations of the
32 problems.

33 Testing multimodal interactions usually requires an additional activity corre-
34 sponding to the presentation of input and output devices to the user. When complex
35 interaction techniques are considered (as in the current application) the presenta-
36 tion of the application to be tested with the user also requires a description of the
37 actual interaction technique. This description goes beyond the typical high-level
38 (task-based) scenarios promoted by usability testing methods (as presented in the
39 previous paragraphs).

40 41 42 **5.5.1.2 Cognitive Walkthrough**

43
44 Cognitive walkthroughs are rarely used in usability testing of multimodal user inter-
45 faces. Quite often, this method is employed in the early phase of the design process,
46 using paper prototypes. With minor adaptations, however, cognitive walkthrough

01 can be successful employed to evaluate multimodal interaction. Following the
 02 method presented in Lewis, Polson & Wharton (1990), the evaluators try to answer
 03 the questions from the cognitive walkthrough evaluation form while conducting the
 04 walkthrough:

- 05 ● Description of the user's immediate goal
- 06 ● First/next atomic action user should take:
 - 07 ● Obvious that action is available? Why/Why not?
 - 08 ● Obvious that action is appropriate to goal? Why/Why not?
 - 09 ● Obvious that action is appropriate to goal? Why/Why not?
 - 10 ● How will user access description of action?
 - 11 ● How will user access description of action?

12 The adaptation needed to assess multimodal 3-D applications involves a careful
 13 preparation of material used to test the learnability of systems. In the case of cog-
 14 nitive walkthroughs, the learnability can be inferred according to the descriptions
 15 provided by users for tasks, actions, and goals. We suggest using the real prototype
 16 installed on a laptop, and giving all participants of the cognitive walkthrough the
 17 ability to try the system before performing the cognitive walkthrough. During the
 18 cognitive walkthrough, the user interface must be projected on the wall, and a paper
 19 version with screen shots of the current task is also used.

20 Important when using the cognitive walkthrough in the evaluation of multimodal
 21 interfaces is the adoption of guidelines to define the questions in the evaluation
 22 form. We believe that the models of the interaction technique (made available in
 23 the formal modeling phase) can support this selection of guidelines (as this will be
 24 shown in the following section dedicated to model-based evaluation).

27 **5.5.2 Model-based Evaluation**

30 While usability evaluation methods are quite efficient for tracking structural usabil-
 31 ity problems (based on ergonomic criteria or navigation problems in an application),
 32 multimodal applications often present fine grain interaction techniques that are dif-
 33 ficult to assess due to their intrinsically complex nature involving concurrent and
 34 time-constrained behaviors.

35 For instance, the description of the temporal evolution presented in Section 5.4
 36 "Modeling the Case Study" shows how complex low-level multimodal interaction
 37 can be. When it comes to testing the usability of such behavior, providing a detailed
 38 description of the behavior to the evaluators is required first, as well as the ability to
 39 to modify such behavior if the results of the usability testing require doing so.

42 **5.5.2.1 Low-Level Scenario Descriptions**

44 Figure 5.11 is called a *marking tree* of a Petri net. It is made up of all the possible
 45 sequences of action through a Petri net. When dealing with Petri net models of
 46 interactive systems, it is quite common that the model is live (i.e., whatever state

AQ2



01 the system is in, there is always a transition, such as a user action, available in
 02 the system). Such so-called *properties* of interactive systems have an impact on the
 03 marking tree that is therefore infinite. While infinite trees are quite difficult to handle
 04 when verification of such systems is considered, this is not a problem for usability
 05 evaluation, because the task is to drive the usage (and evaluation) of the system, and
 06 by definition the task is made up of a finite set of actions. Indeed, when the task is
 07 terminated, the description of other sequences or possible interactions is not relevant
 08 and thus can be ignored.

09 Figure 5.11 shows the set of interaction commands possible on the case study. It
 10 explicitly shows which interaction techniques are available and when (for instance,
 11 from the initial state of the application all actions on the input devices are allowed).
 12 Normal circles represent user actions, while shaded ones represent system reactions
 13 after a user action has occurred (as described in the key on the right-end corner
 14 of Figure 5.11). Arrows connect these actions, making explicit all the possible
 15 sequences of user actions on the input devices. To keep the diagram readable, only
 16 a small part of the interaction space is represented. Indeed, only two releases of
 17 buttons are represented (named “Button Forth Released” in the center and at the
 18 bottom of the figure), while these actions are allowed from any system state (but
 19 depending on the history of the interaction i.e. what the user has previously done).
 20 Similarly the use of speech commands for increasing and reducing rendering speed
 21 (see top left hand side of Figure 5.11) has only been presented from the initial state
 22 but they are available from any state.

23 With respect to the Petri net model of Figure 5.8, it also makes explicit the link
 24 between system reactions and user actions. It is, however, important to understand

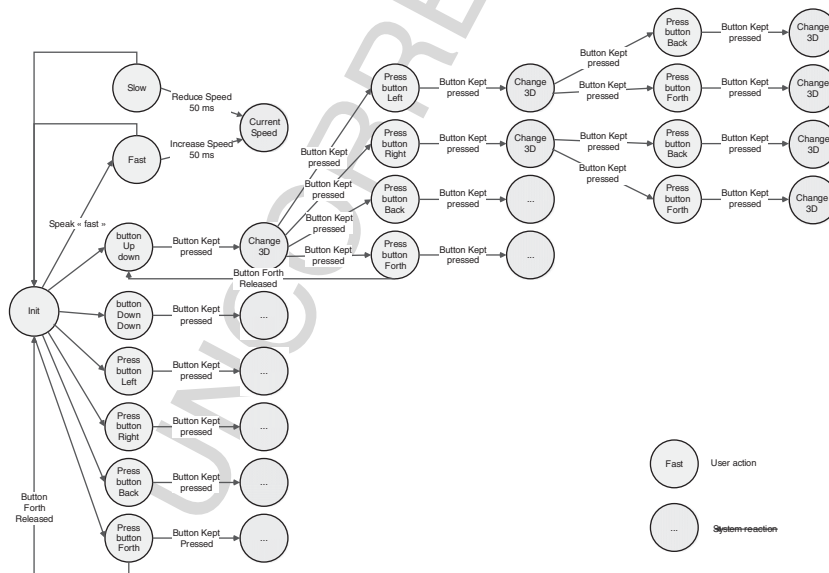


Fig. 5.11 Interaction scenarios from Petri net in Figure 5.8

01 that the interaction space represented in Figure 5.11 is generated from the Petri
02 net model of Figure 5.8 and is not supposed to be constructed manually. While the
03 model of Figure 5.8 is used at run time for driving the execution of the application, as
04 presented in Navarre, Palanque, Dragicevic & Bastide (2006), the interaction space
05 presented in Figure 5.11 is used for describing low-level scenarios. While the Petri
06 net provides an implicit representation of the state space (a given state for the system
07 is described by the distribution of tokens in the places of the net) the marking graph
08 provides an explicit representation of the states (one circle for each state). What is
09 also made explicit in that diagram is that the more the user uses buttons, the fewer
10 options are available. Indeed, the button currently used is not available, and neither is
11 the opposite one (i.e., if the Up button is pressed, neither Up or Down are available).
12 This is why the set of options available to the user reduces very quickly from the
13 quite important one in the initial state.

14 The main contribution of these interaction scenarios is to provide an explicit and
15 complete description of the set of interactions available for the user. This makes it
16 possible to test the usability of an interactive application not only at a high level,
17 such as tasks or goals (as shown with cognitive walkthrough, for instance), but also
18 at a lower level of detail. This makes it possible to test

- 19
20
- 21 • The interaction technique per se (i.e., how difficult it is to manipulate the input
22 devices for low-level tasks like pointing and selecting objects). In the case of
23 multimodal interactions, it is also possible to evaluate the difficulties for users to
24 combine input devices.
 - 25 • The link between the interaction technique and task execution by making explicit
26 what kind of low-level action has to be executed in order to perform higher-level
27 tasks and reach goals.
 - 28 • The complexity of the interaction technique so that users' difficulties in interact-
29 ing with the application can be predicted. To get more figures about this complex-
30 ity, we need to apply our approach to several interaction techniques and to analyze
31 correlations with the results from actual usability tests. We already partially
32 addressed these issues in the domain of Air Traffic Control (Palanque, Bastide
33 & Paterno 1997) and we are now carrying on with more complex interaction
34 techniques and different application domains.
 - 35 • Other complex behaviors related to quantitative temporal behaviors (number
36 of milliseconds for the delays, for instance) can also be represented and thus
37 exploited during usability tests. Indeed, the interaction technique can embed
38 temporal behaviors like the one presented in the case study, and this can have
39 an impact on users' performance. The explicit representation of these temporal
40 evolutions in the interaction technique models makes it possible to incorporate
41 such values while usability experts analyze the results of the usability tests.

42
43 Lastly, this explicit representation of low-level scenarios is useful for selecting
44 the scenarios that will be evaluated with the users, with the explicit purpose of eval-
45 uating comfort and cognitive workload induced by this kind of reduction of the
46 interaction space.

01 **5.5.2.2 Modifying Models to Accommodate Changes**

02
 03 The aim of usability evaluation is not only to identify usability problems in a user
 04 interface, but is also to provide information to the designers to modify the system
 05 and the user interface to reduce or remove the identified problems. Here again,
 06 model-based approaches can be of great interest. According to the usability eval-
 07 uation performed, we describe how some modifications can be incorporated in the
 08 ICO model of Figure 5.8:

- 09
- 10 • Changing the value (increase or decrease) of time related to speech commands—
- 11 this can be done by changing the line delay = old + 50 in the transition slower_,
- 12 for instance, to another amount of increase
- 13 • Changing the maximum speed (increase or decrease) of 3D image rendering—
- 14 this can be done by changing the precondition in transition slower_ or faster_
- 15 to another value than 1000 (maximum) and 100 (minimum)
- 16

17 Other complex behaviors relative to qualitative temporal behaviours can also be
 18 represented and thus exploited during usability tests. For instance, as modeled in
 19 Figure 5.8, all the input modalities are available at all times but another design
 20 choice could have been to allow only the use of a maximum of two input modalities
 21 at a time.

22 In Figure 5.12, we have added a test arc between place *moving* and transition
 23 *faster_*. This means that the voice command *fast* will only have an effect if per-
 24 formed while the 3-D representation of the satellite is currently modified using the
 25 other input devices (the two mice).

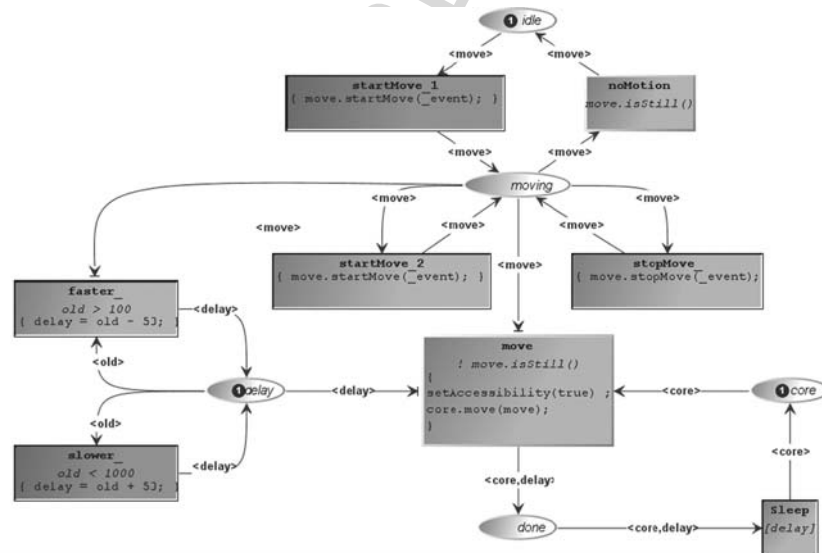


Fig. 5.12 Modification of the availability of a speech command

01 Such modification would have been made explicit to the users by means of the
02 scenarios extracted from the marking graph. This is an example of a modification
03 of a low-level interaction technique that would require usability testing to assess
04 its impact on the overall usability of the system. Similarly, some scenarios could
05 have been selected with the explicit purpose of evaluating comfort and cognitive
06 workload induced by this kind of reduction of the interaction space.
07
08

09 **5.5.3 Standard UEMs and Model-based Evaluation**

10

11
12 During the set-up of a usability test, the characteristics of multi-modal interaction
13 have to be taken into account (see Section 2.2.1. for an extensive discussion):

- 14 ● Which pairs of (device, interaction technique) have to be tested?
- 15 ● How can the user address the system using the various communication channels,
16 and which channel can be used in the various contexts (tasks)?
- 17 ● What types of fission and fusion can be tested (especially in the case of safety-
18 critical systems)?
- 19 ● How can the various dimensions affecting usability evaluation of MMIs be
20 addressed (usage and interpretation of modalities, individual user preferences,
21 context-of-use, and activities supported by the system, etc.)
22

23 When setting up a usability test for multi-modal interfaces, the selection of tasks
24 must be informed by the models. Tasks with high complexity in the low-level mul-
25 timodal interaction must be listed exhaustively. Up to now, this comprehensive list
26 has to be done by the expert describing the tasks to be evaluated. In addition, the
27 frequency of highly complex, low-level interactions has to be estimated, based on
28 the task models. In the case study, the number of synergistic usages of speech and
29 two mice has been counted. Based on this information, the tasks for the usability test
30 representing the low-level interaction are selected. High-level tasks are also selected
31 to conduct the usability test.

32 This way of selecting the tasks for the usability test helps represent all levels of
33 multi-modal interaction. Thus, the results of the usability test are more informative
34 and connected to the precise design of the system.

35 For the set-up of a cognitive walkthrough, the same information about low-level
36 interaction must be used to define the questions during the walkthrough. A result for
37 a specific question might be: “How many modalities can the user cope with, when
38 the user is doing a rotation of a satellite? How many speech commands can the user
39 remember, when he is using additionally two mice, and when he is in a stressed
40 situation?”

41 During the performance of the cognitive walkthrough, the fast adaptability for
42 the multi-modal interaction can be quite useful. For example, during the cognitive
43 walkthrough evaluating the rotation of the satellite using two mice and speech, the
44 idea comes up that an additional back command (speech) would have been helpful
45 to interact with the system. The model of this task can be quickly changed, and the
46 added command can be tested with respect to the learnability of the system.






01 Model-based evaluation helps support standard usability evaluation methods to
02 overcome their reported weaknesses when testing multimodal interfaces. Provided
03 with the adequate tools for editing models and generated marking graphs, the above
04 descriptions might sound easy to follow. Of course, the method will only show its
05 benefits, when the method is carefully set up and conducted.
06
07
08

09 **5.6 Advantages of the Approach and Lessons Learned**

10

11 The main contribution of these interaction scenarios is to make an explicit and
12 complete description of the set of interactions available for the user. This makes
13 it possible to test the usability of an interactive application, not only at a high level
14 like tasks or goals, but also at a lower level of detail. This makes it possible to test
15
16

- 17 • The interaction technique per se (i.e., how difficult it is to manipulate the input
18 devices for low-level tasks like pointing and selecting objects). In the case of
19 multimodal interactions, it is also possible to evaluate the difficulties for users to
20 combine input devices 
 - 21 • The link between the interaction technique and the tasks' execution by making
22 explicit what kind of low-level actions have to be executed to perform higher-
23 level tasks and to reach goals 
 - 24 • The complexity of the interaction technique that is needed to be able to predict
25 users' difficulties in interacting with the application. To get more figures about
26 this complexity, we need to apply our approach to several interaction techniques
27 and analyze correlations with the results from actual usability tests. We already
28 partially addressed these issues in the domain of Air Traffic Control (Palanque,
29 Bastide & Paterno 1997) and we are now carrying on with more complex inter-
30 action techniques and different application domains 
- 31
32

33 Other complex behaviors relative to qualitative temporal behaviors (number of
34 milliseconds for the delays) can also be represented and thus exploited during
35 usability tests. Indeed, the interaction technique can embed temporal behaviors like
36 the one presented in the case study and this can have an impact on the users' per-
37 formance. The explicit representation of these temporal evolutions in the interaction
38 technique models makes it possible to incorporate such intrinsic values while usabil-
39 ity experts analyze the results of the usability tests.

40 Lastly, this explicit representation of low-level scenarios is useful for selecting
41 which scenarios will be evaluated with the users. Indeed, in the field of safety-
42 critical systems, scenario identification is critical as some tasks can be performed
43 very rarely (such as setting a satellite to a survival mode or tasks involving very
44 low probability events like failures). Being able to detect potential difficulties for
45 the users from the analysis of the interaction models can fruitfully influence the way
46 usability tests will be performed.

5.7 Conclusion and Perspectives

In this chapter, we have detailed the intricate problem of usability evaluation of multimodal user interfaces, using a case study on safety-critical systems. Several issues, such as low-level interaction, fusion and fission of events, and complex and temporal behavior, make multimodal user interfaces difficult to evaluate. In particular, the notion of low-level interaction techniques can have a significant impact on the results and thus the interpretation of usability test results. To overcome these difficulties, we have presented a model-based approach for supporting the evaluation of multimodal user interfaces.

Our approach is illustrated via a Space Ground System of satellite control rooms, for which multimodal interaction techniques are fully described by the means of models. More specifically, we have shown how formal models of dialog and low-level interaction can support usability evaluation through systematic, rational, and low-level scenario identification.

We are currently in the phase of performing such model-based evaluation on a real ground segment information treatment system to assess the impact of multimodal interaction techniques on the ease of use and performance. The goal is also to assess the impact of model-based evaluation with respect to more classical usability evaluation techniques for multimodal systems.

The model-based evaluation can be easily conducted in combination with a usability test or a cognitive walkthrough. We want to investigate further combinations of model-based evaluation and other usability evaluation methods in the near future.

It is important to note that we are not claiming that current practices in the field of usability evaluation must involve model-based usability evaluation. Our claim is that in the field of safety-critical interactive systems, and more specifically when multimodal interaction is considered, model-based approaches can support specific activities (like low-level testing scenarios and task identification) that could be otherwise overlooked or not systematically considered. The benefit is higher in that application domain because safety and reliability concerns already call for such formal description techniques. Exploiting the models for usability also reduces the higher development costs of these approaches.

References

- Bach, C., & Scapin, D. (2003). Ergonomic criteria adapted to human virtual environment interaction. In: Rauterberg, M., Menozzi, M. & Wesson, J. (eds.) *IFIP conference on human-computer interaction (INTERACT'2003)* (pp. 880–883), Zurich, Switzerland. IOS Press.
- Baille, L., & Schatz, R. (2005). Exploring Multimodality in the Laboratory and the Field. In Lazzari, G. & Pianesi, P. (Eds.), *ACM international conference on multimodal interfaces (ICMI'2005)* (pp. 100–107). New York: ACM Press.
- Balbo, S., Coutaz, J., & Salber, D. (2003) Towards automatic evaluation of multimodal user interfaces. *Intelligent User Interfaces, Knowledge-Based Systems*, 6(4), pp. 267–274.

- 01 Bastide, R., Navarre, D., Palanque, P., Schyn, A., & Dragicevic, P. (2004). Model-based approach
02 for real-time embedded multimodal systems in military aircrafts. In *ACM International Con-*
03 *ference on Multimodal Interfaces (ICMI'2004)*. New York: ACM Press.
- 04 Bolt, R. A. (1980). Put-that-there: Voice and gesture at the graphics interface. In *Proceedings of the*
05 *7th International Conference on Computer Graphics and Interactive Techniques* (pp. 262–270),
Seattle, USA.
- 06 Bolt, R. E., & Herranz, E. (1992). Two-handed gesture in multi-modal natural dialog. In J. Mackin-
07 lay & M. Green (Eds.), *Symposium on User Interface Software and Technology (UIST'92)*
08 (pp. 7–14). New York: ACM Press.
- 09 Bowman, D., Gabbard, J., & Hix, D. (2002). A survey of usability evaluation in virtual environ-
10 nments: Classification and comparison of methods. *Presence: Teleoperators and Virtual Envi-*
ronments, 11(4), 404–424.
- 11 Brewster, S. A., Wright, P. C., & Edwards, A. D. N. (1994). The design and evaluation of an
12 auditory-enhanced scrollbar. In *Proceedings of the SIGCHI conference on Human factors in*
computing systems (CHI '94) (pp. 173–179). New York: ACM Press.
- 13 Buxton, W., & Myers, B.A. (1986) A study in two-handed input. In: Mantei M., & Orbeton, P.
14 (Eds.), *ACM Conference on Human Factors in Computing Systems (CHI'86)* (pp. 321–326),
15 Boston, Massachusetts. ACM Press.
- 16 Coutaz, J., Nigay, L., Salber, D., Blandford, A., May, J., & Young, R. (1995). Four easy pieces
17 for assessing the usability of multimodal in interaction: the CARE Properties. In S. A. Arnesen
& S. Gilmore (Eds.), *The IFIP Conference on Human-Computer Interaction (INTERACT'95)*
18 (pp. 115–120). Lillehammer, Norway: Chapman & Hall.
- 19 Dillon, R.F., Edey, J.D., & Tombaugh, J.W. (1990). Measuring the true cost of command selection:
20 techniques and results. In J.C. Chew & J. Whiteside (Eds.), *ACM Conference on Human Factors*
in Computing Systems (CHI'90) (pp 19–25), Seattle, Washington. ACM Press.
- 21 Dragicevic P., & Fekete J-D. (2001). Input device selection and interaction configuration with
22 ICON. In A. Blandford, J. Vanderdonck & P. Gray (Eds.), *People and Computer XV – Interac-*
23 *tion without Frontiers: Joint Proceedings of HCI 2001 and IHM 2001* (pp. 448–543). London:
24 Springer.
- 25 Dragicevic, P., Navarre, D., Palanque, P., Schyn, A., & Bastide, R. (2004). Very-high-fidelity
26 prototyping for both presentation and dialogue parts of multimodal interactive systems. In:
27 *DSVIS/EHCI 2004 Joint Conference 11th Workshop on Design Specification and Verification of*
Interactive Systems and Engineering for HCI, Tremsbüttel Castle, Hamburg, Germany, Lecture
28 Notes in Computer Science n° 3425 (pp. 179–199).
- 29 Dybkjær, L., Bensen, N. O., & Minker, W. (2004). New challenges in usability evalua-
30 tion – Beyond task-oriented spoken dialogue systems. In *Proceedings of ICSLP*, (vol. III)
31 (pp. 2261–2264).
- 32 Jöst, M., Haubler, J., Merdes, M., & Malaka, R. (2005). Multimodal interaction for pedes-
33 trians: an evaluation study. In J. Riedl, A. Jameson, D. Billsus & T. Lau (Eds.), *ACM*
International Conference on Intelligent User Interfaces (IUI'2005) (pp. 59–66), San Diego.
34 ACM Press.
- 35 Hinckley, K., Pausch, R., Proffitt, D., & Kassel, N. F. (1998). Two-handed virtual manipulation.
36 *ACM Transactions on Computer-Human Interaction*, 5(3), 260–302.
- 37 Holzapfel, H., Nickler, K., & Stiefelhagen, R. (2004) Implementation and evaluation of a
38 constraint-based multimodal fusion system for speech and 3D pointing gesture. In R. Sharma
& T. Darrell, T. (Eds.), *ACM International Conference on Multimodal Interfaces (ICMI'2004)*
39 (pp. 175–182). New York: ACM Press.
- 40 Kaster, T., Pfeiffer, M., & Bauckhage, C. (2003). Combining speech and haptics for intuitive and
41 efficient navigation through image database. In S. Oviatt (Ed.), *ACM International Conference*
on Multimodal interfaces (ICMI'2003) (pp. 180–187). New York: ACM Press.
- 42 Kabbash, P., Buxton, W., & Sellen, A. (1994). Two-handed input in a compound task. In C. Plaisant
43 (Ed.), *ACM Conference on Human Factors in Computing System (CHI'94)* (pp. 417–423),
44 Boston, Massachusetts. ACM Press.
- 45 Kjeldskov, J., & Stage, J. (2004). New techniques for usability evaluation of mobile systems. *Inter-*
46 *national Journal on Human-Computer Studies*, 60(5), 599–220.

- 01 Lewis, C., Polson, P., & Wharton, R. (1990). Testing a walkthrough methodology for theory-based
02 design of walk-up-and-us interfaces. In J. C. Chew & J. Whiteside (Eds.), *ACM Conference*
03 *on Human Factors in Computing Systems (CHI'90)* (pp. 235–241), Seattle, Washington: ACM
04 Press.
- 05 Memon A., Pollack M., & Soffa M-L. (2000). Automated test oracles for GUIs. *ACM SIGSOFT*
06 *Software Engineering Notes: Proceedings of the 8th ACM SIGSOFT International Symposium*
07 *on Foundations of Software Engineering: twenty-first century applications (SIGSOFT '00/FSE-*
08 *8), Volume 25, Issue 6* (pp. 30–39). ACM Press.
- 09 Navarre D., Palanque P., Dragicevic P., & Bastide R. (2006). An approach integrating two comple-
10 mentary model-based environments for the construction of multimodal interactive applications.
11 *Interacting with Computers, 18*(5), 910–941.
- 12 Navarre, D., Palanque, P., Bastide, R., Schyn, A., Winckler, M., Nedel, L., & Freitas, C.M.D.S.
13 (2005). A formal description of multimodal interaction techniques for immersive virtual reality
14 applications. In M. F. Costabile & F. Paterno (Eds.), *IFIP Conference on Human-Computer*
15 *Interaction (INTERACT'2005)* (pp. 170–183). Berlin: Springer-Verlag.
- 16 Nedel, L., Freitas, C.M.D.S., Jacob, L., & Pimenta, M. (2003). Testing the use of egocentric
17 interactive techniques in immersive virtual environments. In: M. Rauterberg, M. Menozzi
18 & J. Wesson (Eds.), *IFIP Conference on Human-Computer Interaction (INTERACT'2003)*
19 (pp. 471–478), Zurich, Switzerland. IOS Press.
- 20 Nielsen, J., & Mack, R. (1994). *Usability inspection methods*. New York: Wiley.
- 21 Ould, M., Bastide, R., Navarre, D., Palanque, P., Rubio, F., & Schyn, A. (2004). Multimodal and 3D
22 graphic man-machine interfaces to improve operations. In *Proceedings of Eighth International*
23 *Conference on Space Operations* (pp. 435–450), Montréal, Canada.
- 24 Oviatt, S. (1999). Ten myths of multimodal interaction. *Communications of the ACM, Vol. 42*(11),
25 74–81.
- 26 Palanque, P., Bastide, R., & Paterno, F. (1997). Formal specification as a tool for objective assess-
27 ment of safety-critical interactive systems. In *Proceedings of IFIP TC 13 INTERACT'97 con-*
28 *ference* (pp. 323–330), Sydney, Australia, 14–18 July 1997. Chapman & Hall.
- 29 Paternò, F., & Santos, I. (2006). Designing and developing multi-user, multi-device web interfaces.
30 In G. Calvary & J. Vanderdonck (Eds.), *Conference on Computer-Aided Design of User Inter-*
31 *faces (CADUI2006)*. Dordrecht, the Netherlands: Kluwer Academics Publishers.
- 32 Poupirev, I., Weghorst, S., Billingham, M., & Ichikawa, T. (1998). Egocentric object manipulation
33 in virtual environments: empirical evaluation of interaction techniques. In N. Ferreira & M.
34 Göbel (Eds.), *Proceedings of Computer Graphics Forum (EUROGRAPHICS'98)* (pp. 41–52).
35 Malden, MA: Blackwell Publishers.
- 36 Polson, P. G., Lewis, C., Rieman, J., & Wharton, C. (1992). Cognitive walkthroughs: A method
37 for theory-based evaluation of user interfaces. *International Journal of Man-Machine Studies,*
38 *36*(5), 741–773.
- 39 Reason J. (1990). *Human Error*. Cambridge University Press.
- 40 Suhm, B., Myers, B., & Waibel, A. (1999). Model-based and empirical evaluation of multimodal
41 interactive error correction. In M.E. Atwood (Eds.), *ACM Conference on Human Factors in*
42 *Computing Systems (CHI'99)* (pp. 584–591). New York: ACM Press.
- 43 Trevisan, D., Vanderdonck, J., Macq, B., & Raftopoulos, C. (2003). Modeling interaction for
44 image-guided procedures. In *Proceedings of International Conference on Medical Imaging*
45 *SPIE'2003* (pp. 108–118). San Diego: International Society for Optical Engineering (v. 5029).
- 46 Trevisan, D. G., Nedel, L. P., Macq, B., & Vanderdonck, J. (2006). Detecting interaction variables
in a mixed reality system for maxillofacial-guided surgery. In: *SBC Symposium on Virtual Real-*
ity (SVR'2006) (pp. 39–50). Belém do Pará, Brazil: SBC Press.
- Zhai, S., Barton, A.S., & Selker, T. (1997). Improving browsing performance: a study of four
input devices for scrolling and pointing tasks. In S. Howard, J. Hammond & G. Lindgaard
(Eds.), *The IFIP Conference on Human-Computer Interaction (INTERACT'97)* (pp. 286–292).
Sydney: Chapman & Hall.

01
02
03
04
05
06
07
08
09
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

UNCORRECTED PROOF

01 **Chapter 5**

02
03
04
05
06
07
08
09
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46

Query No.	Page No.	Line No.	Query
AQ1	99	41	Please provide Citation
AQ2	111	11	“learnability is not a word - do you mean the ability of the user to learn the system ?
AQ3	115	46	Again, learnability is not a word - do you mean how easily the system can be learned?

UNCORRECTED PROOF