In Minker, W., Bühler, D., & Dybkjær, L. (Eds.) (2004). *Spoken multimodal human-computer dialogue in mobile environments.* Dordrecht: Kluwer Academic Publishers.

Chapter 1

USER MULTITASKING WITH MOBILE MULTIMODAL SYSTEMS

Anthony Jameson
DFKI, German Research Center for Artificial Intelligence
and International University in Germany
jameson@dfki.de

Kerstin Klöckner
DFKI, German Research Center for Artificial Intelligence
kloeckner@dfki.de

Abstract
Users of mobile systems often simultaneously perform some other task, and multimodality tends to give them greater opportunities to do so. One goal in the design of mobile multimodal systems should therefore be the support of effective user multitasking. Previous research in several areas has made many contributions that are relevant to this goal, but some key issues require further work. Using the example of voice dialing with a mobile phone, we discuss task analyses of two voice dialing methods, showing how such analyses can help to identify possible obstacles to the simultaneous performance of voice dialing and other tasks. Detailed observations of users doing multitasking, supplemented with survey results, confirm that these analyses capture important aspects of the multitasking problem; but also that users' decisions and behavior are strongly influenced by factors not covered by the task analyses, such as previous experience and beliefs about social acceptability. Conclusions are drawn concerning the implications of this research for design methods and for future research in support of user multitasking.

Keywords: Multimodal systems, mobile computing, multitasking, task analysis, eye tracking.

1 The Challenge of Multitasking

Mobile interactive systems—for example, handheld and wearable computers, as well as motor vehicle driver interfaces—raise a usability challenge that is encountered to a lesser degree in stationary systems: Users often try to use such a system while simultaneously performing one or more other tasks that are related to their current environment. For example, while performing a *system-related task* such as retrieving information from the web, checking email, or using a navigation system, the mobile user may want to perform an *environment-related task* such as shopping, conversing, or walking down a street. Whether the user switches back and forth between the two tasks or performs both of them concurrently without interruption, we can speak of *user multitasking*.

When a mobile system is also multimodal, the possibilities for user multitasking may be especially appealing: The ability to choose among different input and output modalities for the system-related task may make it easier for a user to perform an environment-related task simultaneously.

This chapter examines the implications of user multitasking for the design of mobile multimodal systems: How can we design such systems so as to ensure that users can successfully engage in the sorts of multitasking that they want to engage in?

1.1 Relevant Research Traditions

A number of research areas have yielded concepts, theories, and empirical results that are relevant to these questions—although, as we will see, some key issues are not yet well understood.

1.1.1 Motor Vehicle Driver Interfaces

An area that has yielded much of the most directly relevant research concerns driver distraction in connection with in-car systems for drivers. Many empirical studies have examined the ways in which the use of such systems can impair driving performance (see, e.g., Green (Green, 2003)). Some research has yielded explicit models of the relationships between driving and the use of in-car interfaces (see, e.g., Wierwille (Wierwille, 1993); Salvucci (Salvucci, 2001)). One way of viewing the goal of the present chapter is as that of extending this type of research to other types of mobile multimodal system, for which the practical consequences of unsuccessful multitasking are not as dramatic—although they can seriously degrade overall usability.

1.1.2 Industrial and Engineering Psychology

In the broader field of industrial and engineering psychology, which subsumes the area just mentioned, designing for multitasking has been the topic of much research involving systems that are typically used when the operator is attending to more than one task. Examples include air traffic control systems and monitoring systems for industrial plants, which may consist of multiple displays. With systems like these, the operator usually receives training and acquires considerable expertise in performing the tasks involved individually and in combination—an advantage not enjoyed by users of many mobile multimodal systems. (Useful summaries of research are offered by Damos (Damos, 1991), and by Wickens and Hollands (Wickens & Hollands, 2000)—see especially chap. 11.)

1.1.3 Wearable Computing

One of the basic motivations underlying the design of wearable computers is the goal of allowing the user to operate them at virtually all times, in particular while performing various other activities (see, e.g., Sawhney and Schmandt (Sawhney & Schmandt, 2000), for a description of a wearable messaging system and of a study of the success with which users were able to integrate it into their daily activities). This goal is reflected in the basic design of the hardware and in the choice of input and output devices. In particular, one general strategy is to choose or invent input and output methods that require little or no visual attention (see, e.g., Brewster et al. (Brewster, Lumsden, Bell, Hall, & Tasker, 2003)). In this field, however, there has so far been less attention to the theoretical and empirical analysis of multitasking than there has been in the first two areas mentioned above.

1.1.4 Multimodal Systems

Similarly, in research on multimodal systems, the goal of supporting multitasking is often mentioned as one benefit of the availability of multiple input and output modalities. Modality Theory (see, e.g., Bernsen (Bernsen, 2001)), a framework for deciding which modalities are appropriate for which purposes, includes some predictions about the suitability of particular modalities for combination with particular types of environment-related activities. But as we will see in this chapter, the mere availability of a method for combining two tasks does not guarantee that users will be able to discover and use it effectively.

1.1.5 Handheld Computing

Handheld computers are not as strongly associated with multitasking as wearable computers are. As Pascoe et al. (Pascoe, Ryan, & Morse, 2000) point out, they are often used in stationary settings (e.g., sitting in a chair), much like laptops. Accordingly, research into their usability has tended to focus more on the consequences of their limited size, capacity, and bandwidth than on their suitability for multitasking. One exception is the work of Pascoe et al. (Pascoe et al., 2000), who studied software for handhelds that was specially designed to be used in conjunction with physically and mentally demanding environment-related tasks like tracking and observing animals in the wild.

1.1.6 Cognitive Psychology

In experimental cognitive psychology, studies of multiple-task performance have a long tradition (see, e.g., Meyer and Kieras (Meyer & Kieras, 1997), and Kieras and Meyer (Kieras & Meyer, 1997), for influential analyses). A typical procedure in such a study is to investigate the ways in which two tasks interfere with each other when they are performed concurrently, with the goal of elucidating the unobservable processing mechanisms involved in one or both tasks. By contrast, when we are designing so as to support multitasking, it may be less important to understand exactly why a particular type of performance decrement arises when two tasks are performed concurrently; we may be more interested in ways of redesigning the system in question so that the decrement does not arise. Despite this difference in overall research goals, many of the theoretical concepts and specific results that have arisen from experimental research have implications for system design.

1.1.7 Human-Computer Interaction

In the field of human-computer interaction as a whole, multitasking has received less attention than it has in the areas mentioned above. The reason may be that until recently almost all of the systems studied were implemented on stationary computers that offered only limited opportunities to combine tasks. Most of the relevant contributions from this area concern support for task switching as opposed to simultaneous uninterrupted execution of tasks (see, e.g., Miyata and Norman (Miyata & Norman, 1986); Cutrell et al. (Cutrell, Czerwinski, & Horvitz, 2001)).

As we will see (e.g., in Sections 1.2 and 3.1), the human-computer interaction field has contributed relevant theories and modeling methods that do not specifically concern multitasking.

1.2 Central Open Issues

Because of some typical properties often shown by mobile multimodal systems, designing for multitasking in this context raises some questions that cannot be answered fully on the basis of the types of research discussed so far.

Issue 1. For any given combination of a system-related and an environment-related task that a user is likely to want to perform simultaneously, how can we ensure that there will exist some suitable method for combining these tasks?

One factor that makes this question difficult is the large number of possible system- and environment-related tasks that a user might want to combine. After all, a mobile multimodal system these days can have as much functionality as a PC of a few years ago; and its user may not be restricted to using it only while driving a car, or only while doing a certain type of field work.

Moreover, as a system becomes more complex, it becomes trickier to ensure that its use will be compatible with the execution of another task, because the number of possible courses that the interaction with the system can take becomes greater.

Issue 2. How can we make sure that the user who wants to combine two tasks can quickly discover a suitable method for doing so?

This issue also follows in part from the relatively large number of possible task combinations, which leads to a large number of situations in which the user is trying to deal with a new task combination that she has never attempted before.

Moreover, designers of a mobile multimodal system will typically offer more than one way of accomplishing each given task, partly in response to Issue 1. But such increased flexibility increases the search space of possible ways of combining the two tasks.

Finally, whereas in industrial and experimental settings users are typically given instruction and/or practice at combining a given pair of tasks, multitasking instruction is not common in other settings. Even when it is offered, it will run up against the general tendency of users to focus immediately on the tasks that they want to perform, as opposed to learning about the system (cf. the *production paradox* discussed by Carroll and Rosson (Carroll & Rosson, 1987)). Even rational deliberation about appropriate methods for performing new tasks seems to be the exception, the rule being the application of methods that have worked in similar settings in the past (cf. Carroll and Rosson's *assimilation paradox*).

Issue 3. What factors that are not directly related to effective task performance influence users' decisions about how to handle a multitasking problem; and how can the designer take these factors into account?

In industrial and experimental settings, users are typically induced to do their multitasking in some tried and tested way that tends to maximize effectiveness and—if applicable—safety. Users who are free to choose methods for performing their tasks may be influenced by more subjective considerations, including those involving social acceptability and personal comfort. Design solutions that do not take such factors into account are likely not to be used in the expected way.

In this chapter, it will not be possible to give comprehensive answers to these three questions—not only for reasons of space, but also because a good deal of further research is required. But we will provide some initial analyses and results so as to make the issues more concrete and to offer glimpses of possible answers. After Section 2 has introduced a suitable example system for analysis, Sections 3 and 4 will offer theoretical discussion based on task analyses. Section 5 will then report on two user studies that shed further light on the issues. Finally Section 6 will summarize the contributions of this chapter to the understanding of the three central issues.

2 Example System

As an especially simple example of a mobile multimodal system, we will consider a cell phone that supports voice dialing. One carefully designed phone of this sort is the Siemens S35i mobile phone. Here are the instructions for voice dialing from the manual of this phone

> In standby: To select, press the lower key on the left side of the phone . . . Then say the name. The phone number will be dialed automatically.

These instructions can be seen as an instantiation of the concept of the *minimal manual* (see, e.g., Carroll et al. (Carroll, Smith-Kerker, Ford, & Mazur-Rimetz, 1987)): They give the user enough information to start performing the task in question; but they do not specify all of the actions that the user will have to perform or exactly what the system will do. They therefore presuppose that the user will be able to discover the details of an appropriate method on her own.

The user will in fact usually find it possible to discover at least one method, since the phone provides information redundantly, in two output modalities, thereby allowing considerable flexibility in the performance of the task. Figure 1 gives an overview of the dialing procedure and the feedback that the phone gives.

As we will see in Section 4, one challenge for the user is to choose a method that fits well with the environment-related tasks that she is performing. So as to be able to understand this problem better, we will now look closely at two different methods for making use of the feedback supplied by the phone.

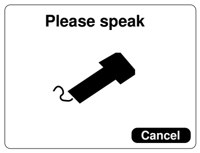
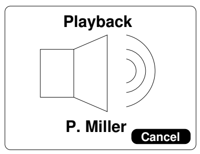
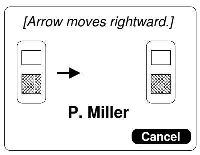
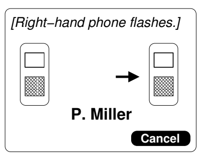
<table>
<tr><th>Display</th><th>Sounds</th><th>User's action</th></tr>
<tr><td>[Default screen]</td><td></td><td>Press button on left–hand side of phone.</td></tr>
<tr><td>
</td><td>Loud beep (audible even when phone is away from ear)</td><td>Say name of callee as previously recorded for voice dialing (e.g., "Peter").</td></tr>
<tr><td>
</td><td>Playback of previously recorded speech sample</td><td>Check name to see if speech recognition was correct.</td></tr>
<tr><td>
</td><td>Melody</td><td>Check that connection is being set up (optional).</td></tr>
<tr><td>
</td><td>Ringing sounds</td><td>Check that callee's phone is ringing (optional).</td></tr>
<tr><td></td><td></td><td>Bring phone to ear and wait for callee to answer.</td></tr>
<tr><td>[Further graphical feedback]</td><td>Ringing sounds; callee's voice</td><td></td></tr>
</table>

Figure 1. *User actions and system feedback during voice dialing with the Siemens S35i mobile phone.*
(Screen display translated from German and redrawn for legibility.)

3 Analyses of Single Tasks

3.1 Eye-Based Dialing

With the method that we will call *eye-based dialing*,[1] the user $\mathcal{U}$ obtains almost all of the necessary feedback from the system via the visual channel.

[1]In connection with speech-based systems, terms such as *eyes-free* are often used. For our analyses, it is clearer to name the primary perceptual channel that is used with a given method.

Accordingly, $\mathcal{U}$ holds the phone in a position that allows her to see its display throughout most of the dialing process, moving it closer to her mouth when speaking into it. The first column in Figure 1 shows a typical sequence of screens that $\mathcal{U}$ might see while voice-dialing the number of Peter Miller and waiting for him to pick up the phone.

To provide a clearer picture of the ways in which the execution of this method can be combined with the performance of other tasks, Figure 2 shows a task analysis of eye-based dialing. The notation used for this and the subsequent task analyses is adapted from the notation that has been used for the CPM-GOMS model (see, e.g., John and Kieras (John & Kieras, 1996)). This notation is in turn based on the notation for PERT diagrams, which are often used in project management. Each box denotes an elementary action that $\mathcal{U}$ has to perform (or, in the case of boxes with dashed borders, optionally can perform). The distribution of the actions over rows indicates the main perceptual, cognitive, or motor resource required by each action.[2] The dimension from left to right is chronological. A line joining two actions signifies that the first action is a prerequisite for the second one. For example, $\mathcal{U}$ has to hold the phone somewhere in front of her eyes in order to be able to read the prompt on the display; and she cannot speak the appropriate form of the name of the callee until she has retrieved it from memory.

Whereas CPM-GOMS aims to provide a fine-grained, quantitative model of parallel processing in expert users, including time estimates for individual operations, the purpose of our simpler diagrams is to support the qualitative analysis of possible conflicts between concurrently executed tasks.

The distinction between *focal* and *ambient* vision (cf. Wickens and Hollands (Wickens & Hollands, 2000), p. 451) will become relevant when we look at the performance of two concurrent tasks (Section 4). In the present single task of eye-based dialing, $\mathcal{U}$ uses only focal vision to read the feedback on the screen.

RECALL STORED SPOKEN NAME is included as an action that makes use of working memory (even though the name is usually retrieved from long-term memory), because it may require a conscious effort of information retrieval or even reasoning on the part of $\mathcal{U}$: As Figure 1 illustrates, the name that $\mathcal{U}$ recorded when setting up the voice dialing for Peter Miller may have been his first name, his last name, or some other variant. Similarly, the action CHECK WRITTEN NAME requires $\mathcal{U}$ to consider whether the written name that appears on the display refers to the intended person—a nontrivial task if the written and spoken forms are very different. (It is in fact important for $\mathcal{U}$ to make this comparison, since the system sometimes misrecognizes a spoken name and begins dialing the number of some other person.)

[2] As is shown in Figure 1, the user actually initiates each task by pressing the button on the left-hand side of the phone; this action is omitted from the task analyses for reasons of space.

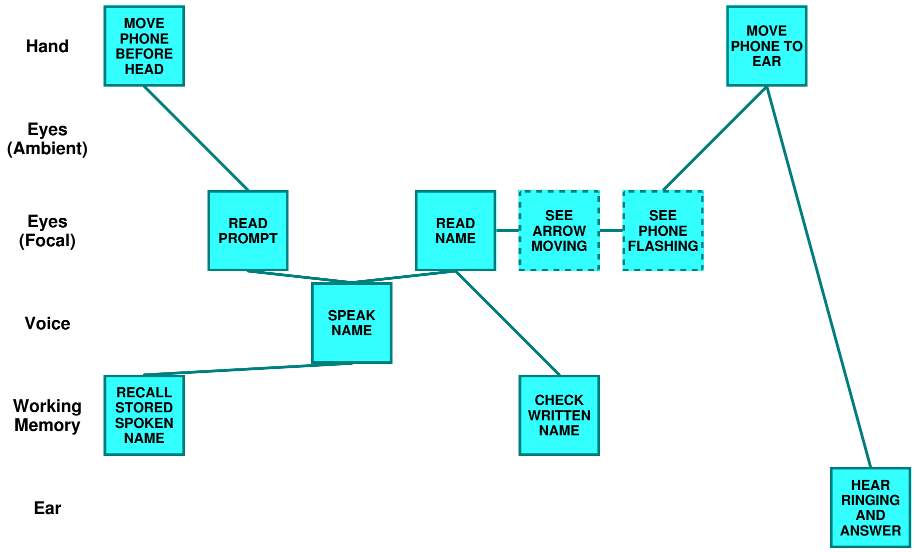


Figure 2. *Analysis of the task of eye-based voice dialing.*
(See the text for an explanation of the notation.)

The actions SEE ARROW MOVING and SEE PHONE FLASHING are optional: $\mathcal{U}$ can refrain from looking at the display while the connection is being established and the phone is ringing; the only risk is that she may bring the phone to her ear after the callee has answered the phone and begun to speak.

3.2 Ear-Based Dialing

Figure 3 describes the method called *ear-based dialing*. Throughout the application of this method, the user can hold the phone in the position that is normally used for talking. This is one respect in which the method is intrinsically simpler than eye-based dialing. The other respect concerns the checking of whether the system recognized the spoken name correctly: The phone's acoustic feedback in the case of correct recognition—the name as $\mathcal{U}$ originally spoke it—is necessarily very similar to the speech that $\mathcal{U}$ has just produced. Therefore, the process of comparing the two names involves no memory retrieval or reasoning. Accordingly, there is no action in this analysis that corresponds to the CHECK WRITTEN NAME of eye-based dialing (Figure 2).

The two methods discussed so far can be combined in various ways. For example, two of the users that we observed (Subjects 5 and 6 in Section 5.1) consistently switched from eye-based to ear-based dialing in the middle of the execution of the task (though at slightly different points). It is possible to create a different graph like the ones in Figures 2 and 3 for each such hybrid method,

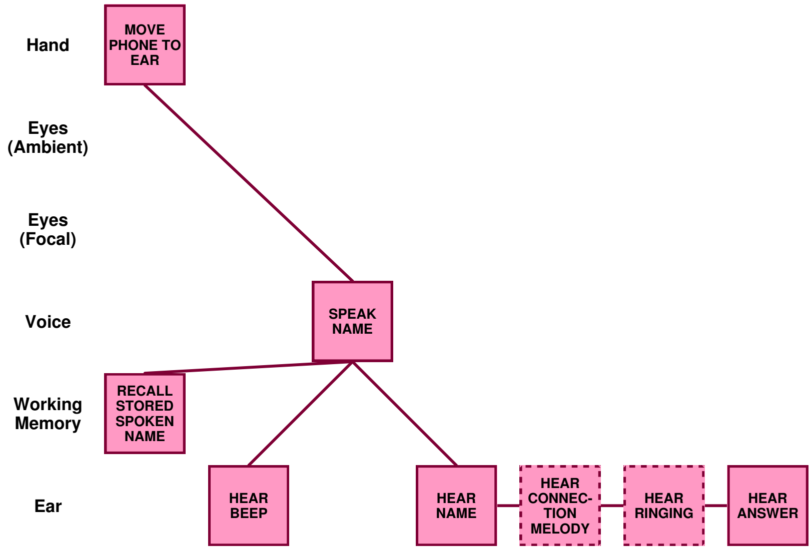


Figure 3. *Analysis of the task of ear-based voice dialing.*

but when discussing such cases we will simply refer to the relevant parts of the analyses for the two pure methods.

3.3 An Environment-Related Task: Walking

The task analyses given so far already suggest some reasons why a subject might consistently prefer eye- or ear-based dialing—for example, a general preference for visual over acoustic feedback, or a desire to avoid the cognitive load imposed by the action CHECK WRITTEN NAME. But in order to analyze possible dependencies of method choice on the nature of an environment-related task, we will have to add an analysis of a typical task that people often like to perform while using a mobile phone. One such task is walking. Since this activity raises interesting issues only when there is some nontrivial navigation for $\mathcal{U}$ to perform, we will consider walking in environments that include obstacles that occasionally require some of the walker's attention.

Figure 4 shows a typical sequence in which $\mathcal{U}$ at first walks straight ahead while keeping an eye out for obstacles, then notices an obstacle, looks at it directly, makes some sort of plan for getting around it, and then executes this plan, continuing to look at the obstacle. (Of course, many other sequences can occur, but this example will be adequate for our purposes.)

The modeling of the use of focal and ambient vision for this task takes into account the fact that a user does not in general have to fixate on an obstacle

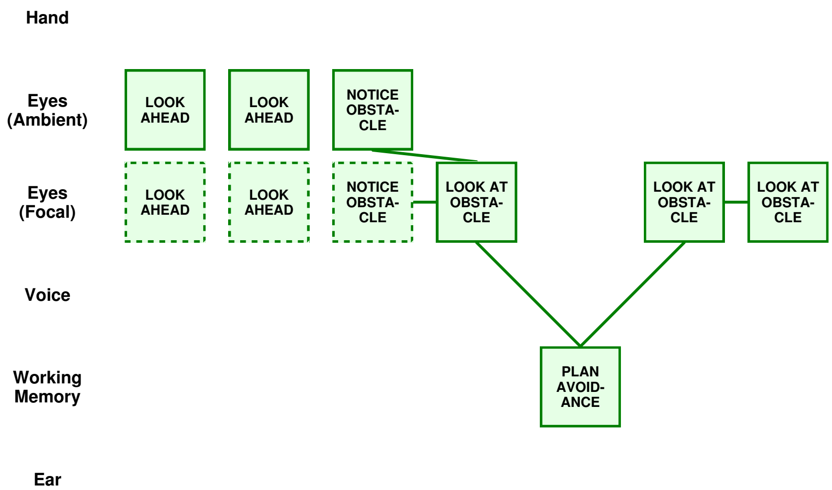


Figure 4. *Simplified analysis of the task of walking through an area that includes obstacles that occasionally demand the walker's attention.*

in order to notice its existence. Instead, using ambient vision to look out for obstacles is considered necessary in this analysis, whereas using focal vision is optional. Only when an obstacle is noticed is $\mathcal{U}$ assumed to fixate on it during the process of avoiding it.

It is actually often possible to navigate around an obstacle entirely without fixating on it, as was confirmed by the eye tracking study described in Section 5.1. Actions related to such easy-to-deal-with obstacles are omitted from the analysis of Figure 4, since they tend not to have much impact on the use of the mobile system.

4 Analyses of Task Combinations

We now turn to the central question of how well users can combine two tasks—in this case, voice dialing and walking. Given that the Siemens phone supports two voice dialing methods, a key question is that of which method users will (and should) choose when walking. When questions like this are discussed—whether in research articles (e.g., Pascoe et al. (Pascoe et al., 2000)) or in nonscientific discussions such as product advertisements—it is customary to reason on an abstract level in terms of the resource demands of entire tasks. For example, several of the student participants in the study summarized below in Section 5.2 made comments like "With ear-based dialing, you can use your

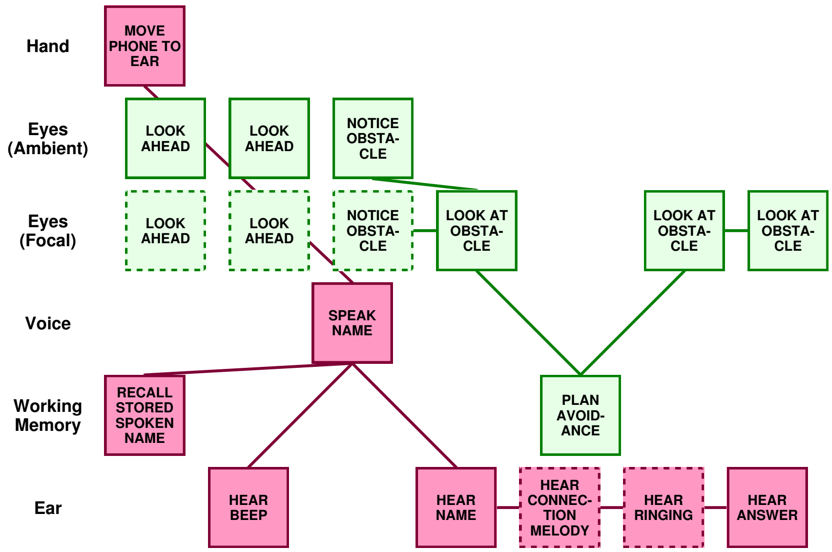


Figure 5. *Superimposition of the task analyses of ear-based dialing and walking, suggesting ways in which the tasks may interfere with each other.*

eyes to watch the street." This type of reasoning is simple enough to be applied with little effort in many situations. But with the help of our task analyses, we will see that the relevant considerations are in general more complex.

4.1 Ear-Based Dialing and Walking

One way to get an initial idea of the problems that can arise when two tasks are combined is simply to superimpose the two task analyses—as is done in Figure 5. The next step is to look for cases where required actions can interfere with each other.

One way of thinking about interference is in terms of *resource conflicts*: cases where $\mathcal{U}$ may face the necessity of simultaneously performing two actions that make use of the same input modality, output modality, or cognitive activity to such an extent that their simultaneous execution is problematic. Various theoretical conceptualizations of resource conflicts have been proposed in the literature, the best-known being Wickens's Multiple-Resource Theory (see, e.g., Wickens (Wickens, 1984)). There is often some question about whether a given example of interference is best explained in terms of resource limitations (cf., e.g., Navon (Navon, 1984); Recarte and Nunes (Recarte & Nunes, 2000)). Still, thinking about task analyses in terms of resources is a useful heuristic method for identifying possible problems, even if they turn out to be explainable in other ways as well.

As an example of a type of interference that clearly does not involve only competition for a limited resource, consider a situation where $\mathcal{U}$'s environment-related task somehow produces beeps that sound like the those produced by the cell phone. $\mathcal{U}$ might have difficulty in telling whether a given beep was relevant to the system-related or to the environment-related task—even if the beeps occurred infrequently and there was no overloading of the auditory channel (cf. Wickens and Hollands (Wickens & Hollands, 2000), p. 454). This sort of interference can be important when the two tasks involved have similar components (e.g., in the case of an air traffic controller who is simultaneously monitoring two airplanes and communicating with their pilots); it appears to be less frequent when a task combination involves a system-related and an environment-related task, since the two tasks are less likely to involve similar stimuli and actions.

With the task combinations that we are considering here, there is no intrinsically necessary relationship between the two tasks' timing; for example, the walking user may encounter an obstacle anywhere in the dialing process.[3] To see where interference might arise, we essentially mentally shift the uppermost analysis horizontally by various degrees and see whether any overlaps can occur between actions in the same row.

When we apply this heuristic approach to Figure 5, the intuitive expectation that few conflicts are likely is largely confirmed.

The main exception concerns the working memory conflict that can arise if $\mathcal{U}$ needs to think about how to deal with an obstacle (PLAN AVOIDANCE) just as she is about to think of the form of the name that she wants to speak (RECALL STORED SPOKEN NAME). If each of these actions is sufficiently complex to place significant demands on the user's working memory, $\mathcal{U}$ may be unable to perform both of them simultaneously without some sort of degradation of performance—for example, stumbling upon the obstacle because of having chosen an inappropriate way of dealing with it; or speaking a form of the name that the system cannot recognize.

As we will see, other task combinations often yield a larger number of instances where actions interfere with each other. But identifying these potential conflicts is just the first step. $\mathcal{U}$ may be able to anticipate a conflict more or less far in advance; and she may be able to adapt her behavior more or less appropriately. In the most favorable case, the potential conflict may have no

[3] Sometimes users combine tasks which are related in content, such as walking around in a city and consulting a handheld navigation system. With such task combinations, relationships between events in the two tasks tend to be more predictable; for example, the walker is relatively likely to consult the navigation system upon reaching a point where she has to choose between two or more directions.

negative consequences. We will look more closely at these possibilities after we have seen more examples of potential conflicts.

This first example illustrates that a potential conflict may not be easy to anticipate on the basis of global task attributes such as being "eyes-free". One reason in the present case is that the conflict concerns working memory demands, which are not tied to any particular input or output channel. A second reason is that the conflict arises only in a fairly special case: when the need to recall the name and to think about an obstacle happen to coincide in time.

More generally, we can see that the variant of CPM-GOMS used here can be used as a tool for identifying possible problems with a system design—but also that the use of this tool is quite different from the application of task analysis notations to single tasks. Because the temporal relationships between the actions that are performed as parts of the system- and environment-related tasks are largely unpredictable, the analyst must consider a much larger number of possible sequences of actions and events; and predictions of global measures such as total execution time or probability of success become less feasible. The number of possible scenarios is further increased when the course of events in the system-related task is itself partly unpredictable—as is the case even with our simple system, as the following analysis will show.

4.2 Unexpected Events

More subtle conflicts can arise when unexpected events occur during the execution of a method. For example, when the system fails to recognize the name spoken by $\mathcal{U}$, instead of hearing the stored name being replayed, the user hears a loud beep; and the display visually prompts the user to restart the voice dialing process. Figure 6 shows the consequences of this event in the case of ear-based dialing. The analysis presupposes (plausibly) that $\mathcal{U}$ does not yet know on the basis of previous experience how she must respond to the beep in order to complete the voice dialing process; and that she must therefore look at the prompt on the display. In other words, the ear-based dialing method turns out to be less eyes-free than $\mathcal{U}$ might expect. If this problem arises at a moment at which $\mathcal{U}$ needs to look at an obstacle, $\mathcal{U}$ will have to figure out quickly how to resolve the conflict. Further difficulties are the natural tendencies of users (a) to respond to an acoustic warning by looking at the display for further information and (b) to carry a task (or subtask) to its completion even when there is a good reason to abort it suddenly—a phenomenon sometimes referred to as an *inertia effect* (see, e.g., Wickens and Hollands (Wickens & Hollands, 2000), p. 445). In the example considered here, these tendencies can cause $\mathcal{U}$ to focus attention on the display, thereby risking a collision with an obstacle.

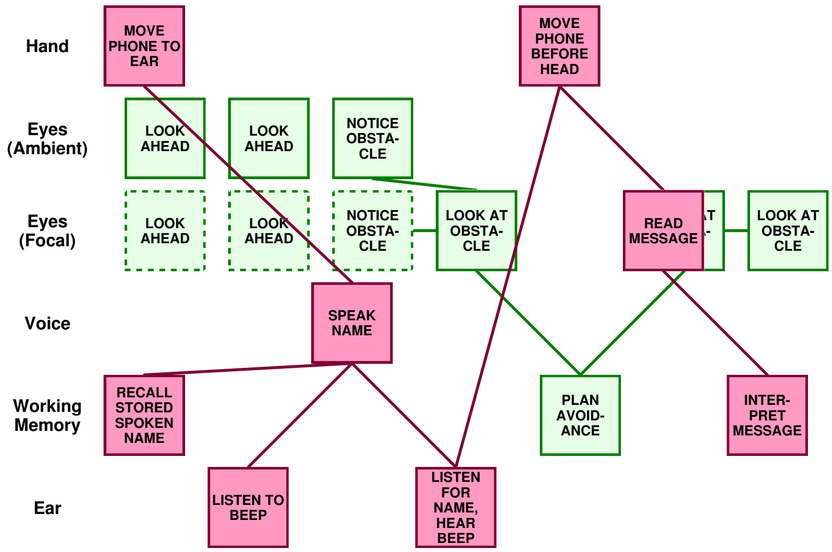


Figure 6. *Superimposition of the analyses of the same tasks as in Figure 5, describing the case where the speech input is not recognized.*

4.3 Eye-Based Dialing and Walking

Now let us consider the voice dialing method that seems less naturally combinable with walking: eye-based dialing (Section 3.1). A rough assessment of the resource demands would suggest that there can be problems in that both this dialing method and walking place fairly heavy demands on the visual input channel. In fact, one might expect that no-one would choose this dialing method in conjunction with walking; but as we will see (Section 5) it is actually fairly popular in this context.

Figure 7 will help us to see (a) why this combination is in fact feasible and (b) what potential problems it gives rise to. (Bear in mind that the temporal relationships among the operators of the two tasks are in general not exactly the ones that arise from the simple superimposing of Figures 2 and 4.) As we can see in the left-hand side of the analysis, an action that requires $\mathcal{U}$ to attend to the display need not lead to a problematic conflict as long as $\mathcal{U}$ is simply looking out for obstacles: $\mathcal{U}$ can then skip the optional fixation of potential obstacles (i.e., the lighter-colored boxes with dashed borders) and rely on ambient vision.

Similarly, as can be seen on the right, if an obstacle requires $\mathcal{U}$ to fixate on it, there need be no serious problem if the only information on the display is the feedback concerning $\mathcal{U}$'s attempt to establish a connection: Since looking at this feedback is optional, $\mathcal{U}$ can look at the obstacle, the only drawback being that she will not know exactly when she needs to move the phone to her ear.

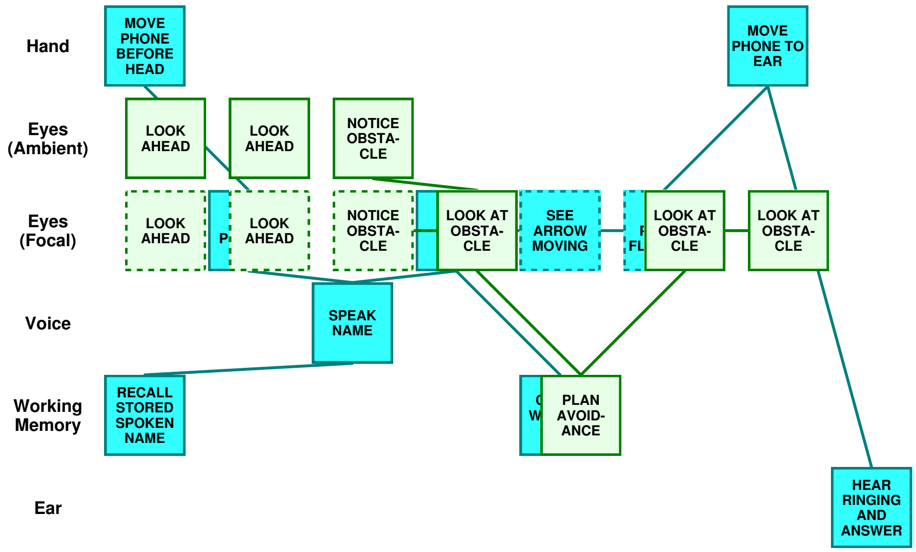


Figure 7. *Superimposition of the task analyses of eye-based dialing and walking.*

On the other hand, the decision to ignore the display in this situation may not be one that $\mathcal{U}$ finds obvious or easy to make:

- It may not be clear to $\mathcal{U}$ that this part of the feedback is less essential than the other parts.
- $\mathcal{U}$ may have established the subgoal of seeing the exact moment at which the callee answers the phone, and she may be reluctant to abandon this subtask.
- The blinking feedback on the display may be visually more salient than the obstacle.
- $\mathcal{U}$ may underestimate the extent to which fixation on the obstacle is necessary.

Probably the most serious conflict involving the visual channel occurs if the need to read the written name (READ NAME) coincides with the need to look at an obstacle (LOOK AT OBSTACLE), as is the case in Figure 7.

- Checking the name is more or less obligatory, since skipping it entails the risk of calling the wrong person.
- Checking the name cannot be postponed, since the callee's phone will start to ring within a few seconds unless $\mathcal{U}$ interrupts the call.
- Looking at the obstacle may be safely postponable only if $\mathcal{U}$ stops walking immediately—something that $\mathcal{U}$ may find difficult.

Finally, an additional working memory conflict can arise because of the need for $\mathcal{U}$ to compare the name written on the display with the name of the intended callee (CHECK WRITTEN NAME, partly obscured in Figure 7 by PLAN AVOIDANCE). Largely the same remarks apply to this conflict as to the similar conflict, involving the operation RECALL STORED SPOKEN NAME, that we already saw in connection with ear-based dialing and walking.

In sum, this approach to combining the two tasks, while basically feasible, gives rise to a number of typical problems and challenges. Before discussing such problems on a more general level, let us look at some data concerning the question: "How do people actually deal with the combination of voice dialing and walking?"

5 Studies With Users

The discussion so far has been based largely on theoretical analysis and general previous knowledge of multitasking and human-computer interaction. In order to ensure that the analyses are in touch with reality, we need to obtain data from users—specifically, to answer the following questions:

- Are our initial task analyses and theoretical considerations consistent with the ways in which users actually perform the tasks in question (individually or in combination)?
  In fact, the results to be summarized below did give rise to a number of refinements of our initial analyses; these refinements were taken into account in the presentation in Sections 3 and 4.
- Within this framework, what general tendencies and preferences do users exhibit?
  Note that the task analyses allow for considerable freedom of user choice—for example, choosing between ear-based and eye-based dialing; choosing which optional operators to execute; and deciding how to prioritize and coordinate the operators of the two tasks. It has long been recognized (see, e.g., Card et al. (Card, Moran, & Newell, 1983), chap. 5) that when a task analysis indicates that users have a choice, different users may exhibit very different behaviors, their choices being determined by a variety of factors.

We will discuss in turn two sorts of user feedback that we have obtained:

1. Fine-grained observation of individual users performing relevant tasks and task combinations.
2. More wide-ranging querying of these users as well as persons whom we have not had the opportunity to observe.

5.1 Detailed Observations

5.1.1 Preliminary Considerations

One traditional way of validating task analyses such as the ones presented above is to (a) use them to derive predictions of task execution times and then (b) compare these predictions with the actually observed times for subjects who perform these tasks (see, e.g., Baber and Mellor (Baber & Mellor, 2001), and Salvucci (Salvucci, 2001), for examples of such analyses involving models of multitasking). Instead of focusing on time predictions, we have chosen to make detailed qualitative observations of users' behavior and compare them with the task analyses, for several reasons:

- The durations of many of the operators that occur in the task analyses (e.g., thinking of the stored name of the person to be called) are not known from previous studies, and they presumably show a good deal of variability from one specific case to the next.
- What we seek is not an overall validation of the models but (a) indications of how they can be refined and (b) information on how users make the many choices that are left open to them.
- The overall design goal with the types of system and task that we are looking at is not so much to minimize execution time as to help users to avoid the many potential problems that can arise with multitasking. Consequently, the ability to predict execution times accurately would not have the practical benefits that it has in some other contexts (e.g., the prediction of the time that telephone operators require to handle calls, as in the classic study of Gray et al. (Gray, John, & Atwood, 1992)).

5.1.2 Method

Recently, various methods have been developed for the collection of detailed data about the behavior of mobile users. One ambitious data collection infrastructure was presented by Oviatt (Oviatt, 2000): With equipment worn by a roaming user, recordings were made of the user's speech and gestural input and of the system's output. Somewhat more recently, Lyons and Starner (Lyons & Starner, 2001) introduced into a similar data collection system a way of capturing the user's view of the current environment: an eyeglasses-mounted video camera that shows the user's field of view. (Other cameras can be added, for example, to record the movements of the user's hands.)

While these methods can yield useful data for the analysis of problems such as those discussed above, what is still lacking is detailed information on what the user is looking at. This information can be obtained with an eye tracker. Eye tracking has a long tradition in studies of driving and the use of in-car systems for drivers (see, e.g., Sodhi et al. (Sodhi et al., 2002)), but handheld

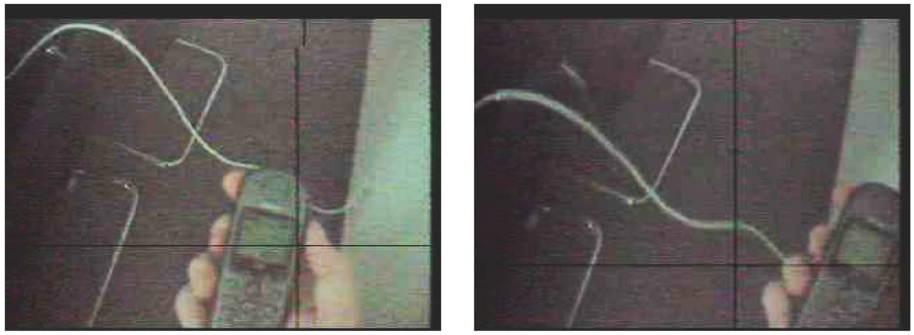


Figure 8. *Typical frames yielded by the mobile eye tracker as Subject 1 performed voice dialing while walking through a room filled with obstacles.* *(The intersection of the two black lines indicates the subject's current point of gaze.)*

and wearable computers raise somewhat different methodological challenges. For our studies, we used an ASL 501 Mobile eye tracker, which is mounted on the subject's head and transmits its data to a stationary control unit via radio waves.

Each of six subjects was first introduced to voice dialing on the Siemens S35i mobile phone. With the experimenter's help, the subject recorded the spoken names of several persons in nearby offices so as to be able to reach them via voice dialing. While the subject was seated at a table, the experimenter introduced the task of voice dialing and demonstrated explicitly that the necessary system feedback could be perceived through either the auditory or the visual modality; but the experimenter gave no instructions or hints about the use of particular modalities. After the mobile eye tracker had been calibrated, the subject was asked to use voice dialing to place a call, (a) while standing still, (b) while walking around a small room filled with a table, some chairs, and various cables, (c) while answering questions asked by the experimenter, and (d) while walking and answering questions simultaneously.[4] In some of the attempts, the phone was actually answered.

When all of the tasks had been performed, the subject watched the video recordings that had been made with the eye tracker. Each recording showed the subject's field of view and direction of gaze during the performance of one task (cf. Figure 8). The subject was asked questions about particular events in the recordings and about his or her general habits and preferences in connection with mobile phones.

[4]These last two task combinations have been analyzed with models such as the ones presented above, but reference to these task analyses must be omitted here for reasons of space.

Later, for each video recording, the experimenter listed the recorded actions and events along with the time intervals during which they occurred.

5.1.3 Subjects 1, 2, and 3: Eye-Based Dialing

Subjects 1, 2, and 3 used the eye-based dialing method for all task combinations.[5] Figure 8 shows two images that indicate the focus of Subject 1's attention as he dialed while walking around the room. Overall, he showed a strong and consistent tendency to fixate on the display, using at most ambient vision to obtain visual feedback from other sources, including information about obstacles. He only occasionally glanced away from the display briefly, apparently to look at an obstacle such as an electric cable.

The interview revealed that this tendency is by no means the result of lack of experience with such tasks: This subject had owned exactly this model of mobile phone for more than 1.5 years, and he reported having used the voice dialing functionality frequently, always with eye-based dialing. When asked about its use during driving, he admitted that he sometimes held the phone in front of his eyes with one hand while steering with the other hand.

When asked why he did not sometimes use ear-based dialing, he said that the acoustic feedback "got on his nerves", because it was "too loud". On the whole, he characterized the ear-based method as "stupid". He admitted that he had sometimes walked into objects while performing eye-based dialing, but he said that he had gotten used to doing so and therefore no longer worried about it.

Subject 2 likewise showed a general preference for eye-based dialing, but she was less strongly inclined to look at the display continually. While walking, she sometimes glanced away from the display to look at obstacles. Like Subject 1, she explained her general preference for eye-based dialing in terms of a consideration that is not represented in the theoretical analyses presented above (but which has been mentioned by several other subjects): It is not customary to hold a mobile phone to one's ear while dialing a number manually, because it is not possible to depress the keys accurately in this position. Subject 2 felt influenced by this habit, even though of course voice dialing is possible in the phone-to-ear position.

Subject 2 also creatively extended the method of eye-based dialing by making some use of the acoustic feedback even while holding the phone in front of her: She said that she was able to hear the playback of the recorded name faintly, thereby being able to benefit from its directness and simplicity while still looking at the display. In fact, further exploration shows that the phone

[5]For clarity of exposition, we discuss the results for the six subjects in a different order than the order in which they were observed.

can be usefully held in a large number of positions between the two extremes considered so far (at the ear vs. far enough in front of the eyes to allow easy focusing on the display). Each intermediate position yields a different point in the tradeoff between the legibility of the information on the display and ability to hear the details of the acoustic feedback. This example illustrates a general challenge for the analysis and prediction of users' behavior where multitasking is concerned: Users cannot be counted upon to restrict themselves to the limited set of elementary actions that are explicitly supported by the system. When they try to operate the system while performing some environment-related task that was not specifically (or successfully) taken into account in the design of the system, they may be inclined to distort the usual actions—or invent new ones—in a creative attempt to fulfill the set of constraints posed by the task combination. Moreover, the number of possible action variants of this sort may be large: Whereas the user of a conventional stationary application like a spreadsheet can hardly perform any actions other than the ones foreseen by the designer (e.g., looking at the screen, selecting menu items, and pressing keys in the normal way), mobile multimodal systems can be held in a variety of positions relative to the user's body and to the environment, and they can be operated in a continuum of ways (e.g., the user's speech can have different speeds, volumes, and styles of articulation). Accordingly, any set of task analyses such as the ones presented above should in general be viewed as representing a sample of the possible methods; and the analyst should be willing to include actions and methods that do not necessarily correspond to the designer's intentions.

Subject 3 likewise used eye-based dialing consistently. He explained later that he was skeptical about the accuracy of the speech recognition. In terms of the task analyses of ear-based dialing presented in Section 4, this subject appeared to believe that the course of events shown in Figure 6, in which the user has to deal with a recognition failure, was sufficiently frequent to constitute a reason to avoid ear-based dialing.

This subject avoided possible interference between walking and eye-based dialing simply by beginning to walk only after he had spoken the name and confirmed that a connection had been established.

5.1.4 Subject 4: Ear-Based Dialing

Whereas the results described so far would suggest a general preference for eye-based dialing for all task combinations, Subject 4 showed the opposite tendency: He used ear-based dialing for all task combinations—even those that involved the environment-related task of conversing with the experimenter, which produces a good deal of auditory input. Two of the reasons that he gave for this preference during the interview correspond to considerations taken into account in our analysis of ear-based dialing (Figure 3):

1. He noticed that the acoustic confirmation of the callee's name is simpler and more natural to interpret than the printed confirmation.
2. He reasoned that ear-based dialing involves less movement of the arm, since the phone is held in just one position throughout the procedure.

Note that these two points concern only the properties of ear-based dialing itself, not its relationships with the methods for environment-related tasks.

Another reason that Subject 4 gave was similar to Subject 2's reference to previous experience, although it referred to a different experience: Subject 4 said that it had never occurred to him before to hold a phone in front of his eyes while speaking into it, since he had never used voice dialing before.

5.1.5 Subjects 5 and 6: Hybrid Methods

Subject 5 quickly learned to use eye-based dialing for the first part of the task and ear-based dialing for the second part: While speaking the name, he always looked at the phone. During the performance of his first task, he continued to look at the display after speaking the name, but he was irritated by the discrepancy between the spoken name ("Sebastian") and the printed name ("Müller"). In subsequent tasks, to avoid further confusion, he moved the phone to his ear immediately after having spoken the name. That is, his reason for switching to ear-based dialing was the same as the first reason given by Subject 4; but instead of using ear-based dialing throughout, he continued to begin the task with the eye-based method.

This subject (like Subject 3) avoided interference between walking and the eye-based part of his method by either standing still or walking slowly while executing this part of the method. Similarly, during the ear-based part of the method, if the spoken name was not recognized at all, the subject immediately stopped walking and looked back at the display (cf. Figure 6).

Subject 6 likewise used a hybrid combination of eye-based and ear-based dialing. Unlike Subject 5, he continued to look at the phone until the recognition of the name had been confirmed; only then did move the phone to his ear (and begin walking, if walking was required).

5.1.6 Conclusions From the Detailed Observations

Although we cannot generalize confidently on the basis of the observation of six subjects, it is striking that each subject consistently used one method (whether eye-based, ear-based, or hybrid), regardless of whether they were walking, conversing, or standing still. Some of the reasons for their method choices can be understood in terms of the task analyses, while others involve previous experience or subjective factors; but considerations related to the relationships between voice dialing and the environment-related tasks evidently did not influence their choice of a method.

Although each subject stuck to one method for all environment-related tasks, their behavior did in some respects take into account the possible interferences between the system-related and the environment-related task. For example, when the environment-related task was walking, they would sometimes slow down or stand still while performing a complex part of the dialing task. The general strategy of suspending execution of one task while dealing with a demanding part of another task has great generality, and subjects have presumably applied it in countless previous situations. By contrast, choosing the most suitable modality for a particular task combination is a more novel problem, whose solution cannot make such direct use of previous experience. Hence the ways in which our subjects took into account the demands of multitasking are consistent with the previous research on learning and method choice mentioned above in Section 1.2.

An alternative to the type of observation performed here would be more longitudinal observation during which users were given numerous opportunities to deal with each specific task combination. It would be interesting to see to what extent method choices became more sensitive to the nature of the environment-related tasks. But the brief observations reported here are actually more relevant than longitudinal studies to the many situations in which users have to deal with a novel combination of system- and environment-related tasks.

5.2 Survey of a Larger Sample of Potential Users

To get a broader picture than is possible with the necessarily limited number of subjects that we can observe in detail, we elicited questionnaire responses from 22 technically sophisticated students from the University of Applied Sciences of Zweibrücken, Germany. Although 21 of the respondents owned a mobile phone, only 1 of these had used any form of voice dialing regularly, and only 3 had done so even occasionally.

The two methods of voice dialing were demonstrated to these respondents with the help of projected slides. They were asked to state on a questionnaire which method they would use in each of several situations and to write down their principal reasons.

The results concerning the choices (in Table 1.1) suggest a sensitivity to the demands of the environment-related tasks that is generally consistent with the sort of task analysis that we have presented above—but inconsistent with the inflexibility of method choice that was shown by the six subjects whom we observed. This discrepancy can be understood in terms of the difference between the demands placed on subjects by a questionnaire study and real task performance, respectively: The questionnaire study may encourage subjects to

Table 1.1. *Expressed Preferences for Eye-Based or Ear-Based Dialing Among 22 Questionnaire Respondents.*

Task Combination	Eye−Based Dialing	Ear−Based Dialing	No Overall Preference	Would Not (Voice−)Dial at All
Sitting still	10	7	5	
Walking down a street	5	17		
Walking down stairs	2	14		6
Driving a car		22		
Conversing with a friend	16	4		2

Note. Each entry shows the number of respondents who gave the response in question.

indicate a preference for different methods for different task combinations and to deliberate rationally about the considerations underlying their choices; subjects actually using the system are likely to be more concerned about performing their tasks successfully.

Although the responses concerning method choice should not be seen as reliable predictors of actual behavior, they do give us a picture of the relevant knowledge and ideas that people have—especially when supplemented with the reasons that the respondents gave for their choices. Since the reasons given overlap considerably with those mentioned by the six subjects that we observed, an overview covering both groups is given in Table 1.2. (Considerations that apply only when the environment-related task is walking down stairs, driving, or conversing, are omitted from the table, because the relevant task analyses have not been presented in this chapter.) Respondents' comments sometimes suggest new hypotheses about factors that may influence behavior, which can be checked through observation of behavior in systematically varied conditions. Also, perhaps users' understanding of possible types of interference among tasks can be leveraged in training and instruction, even if they do not apply it spontaneously.

It can be seen that several of the reasons mentioned correspond to considerations that can be derived from the task analyses presented earlier in this chapter. An approximately equal number of reasons, however, concern different factors. The implications of these results will be discussed in the final section.

Table 1.2. *Overview of Reasons Mentioned by 28 Subjects for Preferring Eye-Based or Ear-Based Dialing.*

Reasons for Preferring Eye−Based Dialing

In General	In Combination With Walking
Reasons related to task analyses:	
The information on the display is less ambiguous, and it's easier to understand.	I think I can see the obstacles well enough just with peripheral vision.
I have to look at the display anyway if the spoken name is not recognized.	
Reasons not related to task analyses:	
The beeps are funny / too loud; they get on my nerves	There's too much noise on the street for me to hear the acoustic feedback.
I find it uncomfortable / silly to hold my arm up for several seconds while waiting for a connection.	
I'm not used to holding a mobile phone to my ear while dialing a number.	

Reasons for Preferring Ear−Based Dialing

In General	In Combination With Walking
Reasons related to task analyses:	
As soon as the person I'm calling picks up the phone, I have to bring the phone to my ear anyway.	You need your eyes in order to be able to look at the street and the environment.
I find it more natural to have the name confirmed by speech, because I entered it by speech in the first place; and the printed name may be entirely different anyway.	
Reasons not related to task analyses:	
I'm used to holding a phone to my ear when speaking into it.	

6 The Central Issues Revisited

By way of summary, let us look back at what we have learned about the three central issues introduced in Section 1.2.

Issue 1: For any given combination of a system-related and an environment-related task that a user is likely to want to perform simultaneously, how can we ensure that there will exist some suitable method for combining these tasks?

The discussion of the task analyses in Section 4 showed that a simplified version of the CPM-GOMS notation can help us to find potential problems with a method for the handling of a combination of tasks—even if the system's design explicitly took that particular combination into account, as is the case with the task of walking, which is quite well supported by ear-based dialing. Note that some of the potential problems uncovered in this way actually occur with only a low frequency (e.g., when two independent events happen to coincide in time, or when some exceptional individual event occurs). That is, this sort of problem can escape the attention of even the diligent empirical researcher and the experienced user.

In some cases the potential problem could be eliminated through a relatively minor design improvement (e.g., eliminating system prompts that the user has to respond to within a certain time window); in other cases, a solution would require the user to adopt a somewhat different method—or an appropriate task-switching strategy.

Issue 2: How can we make sure that the user who wants to combine two tasks can quickly discover a suitable method for doing so?

Improvements to the methods offered by the system will in general also make it easier for users to discover a suitable method for a given situation: If the available methods are largely free of hard-to-anticipate traps and drawbacks, the challenge of finding a suitable method is more surmountable.

But even if, say, the methods of eye-based and ear-based dialing were optimized, we could not assume that users would consistently choose the more appropriate method. The subjects in our studies showed some intuitive awareness of the properties of methods that are relevant to multitasking, but their understanding fell far short of a full grasp of the relevant considerations—as indeed we would expect. Their decisions about which method to use in a given situation were influenced by previous experience with similar phones and by the subjective factors discussed in connection with Issue 3.

The question of how to encourage appropriate method choice is too complex for thorough treatment in this chapter, and it requires further research in connection with mobile multimodal systems. A number of relevant more general strategies were already proposed in the seminal paper by Carroll and Rosson (Carroll & Rosson, 1987). A thorough discussion with regard to a very different type of system (computer-aided design systems) can be found in recent works of Bhavnani and colleagues (see, e.g., Bhavnani and John (Bhavnani & John, 2000)). The most obvious idea is to provide a certain amount of instruction about method choice (e.g., "When you are voice dialing while walking, it is usually best to hold the phone to your ear throughout the process"). Another strategy is simply to limit the number of available methods: For example, a single method that is at least minimally suitable in all situations may be preferable to a set of methods that includes an optimal method for each situation but which also makes it likely that the user will choose methods that are poorly suited to the current situation.

Issue 3: What factors that are not directly related to effective task performance influence users' decisions about how to handle a multitasking problem; and how can the designer take these factors into account?

One type of subjective factor that emerged in our interviews and survey results concerned social acceptability considerations, for example, the notion that a certain way of holding a phone looks silly. It has long been recognized that considerations of this type are especially important with wearable and handheld computers, because these are often used in a variety of social contexts. We see that these acceptability constraints concern not only the basic appearance of a device (e.g., the conspicuousness of a head-mounted display) but also the specific ways in which a device is used—which may be evaluated differently by different users.

Other subjective considerations are related to the user's physical comfort, such as the ideas that the beep tones are too loud or that the position in which the phone must be held is uncomfortable. Such properties may fall outside of the acceptable range especially frequently when mobile multimodal systems are involved. For example, it is harder to find a single optimal beep volume when a device is used in a wide range of physical environments; and allowing users to control these aspects of the interaction can be problematic because of the relatively small communication bandwidth and the frequency with which changes would be required. Therefore, designers should consider the possibility that a basically suitable method for handling a given task in a given context will sometimes be found unacceptable for reasons involving physical comfort and therefore be rejected by the user in favor of an intrinsically less suitable method.

A great deal of further work is required on the questions raised in this chapter. We hope to have convinced the reader of their importance, provided some initial methods and results, and given some concrete ideas about the next steps to be taken.

Acknowledgments

This paper was prepared in the contexts of the Evaluation Center for Language Technology Systems of the DFKI project COLLATE, which is funded by the German Ministry of Education, Research, Science, and Technology (BMB+F) and the project READY of Saarland University's Collaborative Research Center 378 on Resource-Adaptive Cognitive Processes, which is funded by the German Science Foundation (DFG). Marie Norlien helped to develop the methodology employed in the eye tracking studies. We thank the 22 students of the University of Applied Sciences of Zweibrücken, who participated as subjects in the study described in Section 5.2.

References

Baber, C., & Mellor, B. (2001). Using critical path analysis to model multimodal human-computer interaction. *International Journal of Human-Computer Studies*, *54*, 613–636.

Bernsen, N. O. (2001). Multimodality in language and speech systems—from theory to design support tool. In B. Granström (Ed.), *Multimodality in language and speech systems* (pp. 93–147). Dordrecht: Kluwer.

Bhavnani, S. K., & John, B. E. (2000). The strategic use of complex computer systems. *Human-Computer Interaction*, *1*(2/3), 107–137.

Brewster, S., Lumsden, J., Bell, M., Hall, M., & Tasker, S. (2003). Multimodal 'eyes-free' interaction techniques for wearable devices. In L. Terveen, D. Wixon, E. Comstock, & A. Sasse (Eds.), *Human factors in computing systems: Chi 2003 conference proceedings* (pp. 473–480). New York: ACM.

Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction.* Hillsdale, NJ: Erlbaum.

Carroll, J. M., & Rosson, M. B. (1987). The paradox of the active user. In J. M. Carroll (Ed.), *Interfacing thought: Cognitive aspects of human-computer interaction* (pp. 80–111). Cambridge, MA: MIT Press.

Carroll, J. M., Smith-Kerker, P. L., Ford, J. R., & Mazur-Rimetz, S. A. (1987). The minimal manual. *Human-Computer Interaction*, *3*, 123–153.

Cutrell, E., Czerwinski, M., & Horvitz, E. (2001). Notification, disruption and memory: Effects of messaging interruptions on memory and performance. In *Proceedings of Interact 2001* (pp. 263–269). Tokyo.

Damos, D. L. (Ed.). (1991). *Multiple-task performance.* London: Taylor and Francis.

Gray, W. D., John, B. E., & Atwood, M. E. (1992). The precis of project ernestine or an overview of a validation of goms. In P. Bauersfeld, J. Bennett, & G. Lynch (Eds.), *Human factors in computing systems: Chi '92 conference proceedings* (pp. 307–312). New York: ACM.

Green, P. (2003). Motor vehicle driver interfaces. In J. A. Jacko & A. Sears (Eds.), *Human-computer interaction handbook* (pp. 844–860). Mahwah, NJ: Erlbaum.

John, B. E., & Kieras, D. E. (1996). The goms family of user interface analysis techniques: Comparison and contrast. *ACM Transactions on Computer-Human Interaction*, *3*, 320–351.

Kieras, D. E., & Meyer, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 2. accounts of psychological refractory-period phenomena. *Psychological Review*, *104*(4), 749–791.

Lyons, K., & Starner, T. (2001). Mobile capture for wearable computer usability testing. In T. Martin & V. Bahl (Eds.), *Proceedings of the Fifth International Symposium on Wearable Computers* (pp. 69–76). Los Alamitos, CA: IEEE Computer.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. basic mechanisms. *Psychological Review*, *104*(1), 3–65.

Miyata, Y., & Norman, D. (1986). Psychological issues in support of multiple activities. In D. A. Norman & S. W. Draper (Eds.), *User centered system design: New perspectives on human-computer interaction* (pp. 265–284). Hillsdale, NJ: Erlbaum.

Navon, D. (1984). Resources—a theoretical soup stone? *Psychological Review*, *91*(2), 216–234.

Oviatt, S. (2000). Multimodal system processing in mobile environments. In *Proceedings of the 13th Annual ACM Symposium on User Interface Software and Technology* (pp. 21–30). San Diego, CA.

Pascoe, J., Ryan, N., & Morse, D. (2000). Using while moving: Hci issues in fieldwork environments. *ACM Transactions on Computer-Human Interaction*, *7*(3), 417–437.

Recarte, M. A., & Nunes, L. M. (2000). Effects of verbal and spatial-imagery tasks on eye fixations while driving. *Journal of Experimental Psychology: Applied*, *6*(1), 31–43.

Salvucci, D. D. (2001). Predicting the effects of in-car interface use on driver performance: An integrated model approach. *International Journal of Human-Computer Studies*, *55*, 85–107.

Sawhney, N., & Schmandt, C. (2000). Nomadic radio: Speech and audio interaction for contextual messaging in nomadic environments. *ACM Transactions on Computer-Human Interaction*, *7*, 353–383.

Sodhi, M., Reimer, B., Cohen, J. L., Vastenburg, E., Kaars, R., & Kirschenbaum, S. (2002). On-road driver eye movement tracking using head-mounted devices. In *Proceedings of ETRA 2002: Eye Tracking Research and Applications Symposium* (pp. 61–68). New Orleans, LA.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 63–102). Orlando, FL: Academic Press.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Wierwille, W. W. (1993). Visual and manual demands of in-car controls and displays. In B. Peacock & W. Karwowski (Eds.), *Automotive ergonomics* (pp. 299–320). London: Taylor and Francis.