3 THE AUDIO DESKTOP

3.1 INTRODUCTION

The phrase desktop no longer conjures up the image of a polished high-quality wooden surface. The pervasiveness of computing in the workplace during the last decade has led to the concept of a virtual electronic desktop—a logical workspace made up of the documents one works with and the applications used to operate on these documents. Progressive innovations in the Graphical User Interface (GUI) have helped strengthen this metaphor—today, the typical desktop enables the user to organize the tools of his trade by dragging and dropping graphical icons into a visual two-dimensional workspace represented on the computer monitor. Given this tight association between visual interaction and today’s electronic desktop, the phrase audio desktop is likely to raise a few eyebrows (Or should it be earlobes)?

This chapter focuses on the workspace in an auditory environment and defines the audio desktop in terms of the features of auditory displays introduced in Sec. 2.5.3. Using the speech-enabling approach outlined in Sec. 1.5, we demonstrate how the functionality of the electronic desktop can be exposed through an auditory interface. The attempt is not to speak the visual desktop; rather, we identify the key user-level functionality enabled by the modern electronic desktop and describe in detail how this
can be translated to an auditory environment. In visual interaction, the user actively browses different portions of a relatively static two-dimensional display to locate and manipulate objects of interest. Contrast this with auditory displays that are characterized by the temporal nature of aural interaction; here, the display—a one-dimensional stream of auditory output—scrolls continuously past a passive listener. This disparity between aural and visual interaction influences the organizational paradigms that are effective in auditory interaction. The purpose of this chapter is to systematically investigate the design of an effective audio desktop; later chapters demonstrate these techniques in terms of a concrete implementation. The steps in evolving such an audio desktop can be enumerated as:

- Identify user functionality enabled by the electronic desktop,
- Exploit features of auditory displays to enable equivalent functionality and
- Evolve organizational paradigms for aural interaction that compensate for the temporal, one-dimensional nature of audio by exploiting other features of aural interaction.

In Sec. 3.2, we outline the underlying abstractions that make up today’s visual interfaces. Features of today’s electronic desktop and the key user level functionality they enable are enumerated without reference to specific GUI implementations. Dialogues constitute a key component of all interactive interfaces. In Sec. 3.3, we define conversational gestures—the basic building blocks used in constructing complex dialogues, and illustrate each conversational gesture with examples from today’s graphical interfaces. Finally, in Sec. 3.4, we develop equivalent abstractions for the audio desktop.

### 3.2 THE VISUAL DESKTOP

This section enumerates the features of the visual electronic desktop and the user-level functionalities these enable. The discussion does not restrict itself to any single GUI implementation from the past such as Windows 95. The goal is to enumerate the best features of a visual desktop in order to motivate the subsequent sections on the design of an audio desktop.

#### 3.2.1 Windows, Icons, Menus And Pointers

Traditionally, computer interfaces have been made up of a visual display for conveying output and a keyboard for communicating user input. This has been augmented over the years by a pointing device as an additional means of expressing user intent. The visual desktop as realized in all GUI environments today is characterized by four primary abstractions:

**Windows** The visual real estate available on a computer display is divided into rectangular regions for displaying logically separate items of information. Multiple