Chapter 3
Multimodal User Interface in E–Learning Environments: A Proposed Architecture

André Constantino da Silva
UNICAMP, Brazil & IFSP – Campus Hortolândia, Brazil

Heloísa Vieira da Rocha
UNICAMP, Brazil

ABSTRACT
Multimodal interaction is a proposal to turn the interaction between humans and machines more natural, increasing the usability, flexibility, and convenience of one application. Improve an application with multimodal features impacts on its architecture and, to describe the main components to treat the multimodality, some architecture models are proposed in the literature, including for Web multimodal systems. E-Learning environments are Web-based systems and need a good usability, flexibility, and convenience: requirements that can be improved with implementation of multimodal features on them. Since they have their own peculiarities, we need a more specific multimodal architecture model described in such a way to reuse the components built for multimodal systems and to connect them with the e-learning environment components. This chapter proposes an architecture for multimodal e-learning environments. A viability study was done in the Ae, an e-learning environment developed using a component-based development process, with components to treat the pen and touch modalities.

INTRODUCTION
Devices, such as smartphones and tablets, are becoming increasingly popular; most of them have touch screen displays, access to the Internet, and enough computing power to process Web pages. The number of input hardware on conventional computers (desktop and laptops) is increasing, with computing devices equipped with touchscreen devices, microphones, and cameras. Most of the computational devices have an output display, but the size can vary widely. But others output data is possible as voice synthesis and tactile feedback.

One of the mobile devices that are gained repercussion is the tablet, a small and easy-to-carry computational device with a screen between 7 to
Multimodal User Interface in E-Learning Environments

10 inches in which users interact using fingers. Another device is the Tablet PC (Tablet Personal Computer), a device with height as similar to a notebook and had an input device similar a pen, called stylus. The paper and pen metaphor implies that tasks performed before in paper, like draw or manuscript writing, can be more natural in the Tablet PC than in another computing devices. Some Tablet PCs have touch and pen-sensitive screens, allowing users interact using pen or fingers. Resuming, the Tablet PC has the following hardware characteristics: (i) Pen-sensitive screen; (ii) Screen that allow different positions; (iii) Wireless network access by WLAN (Wireless Local Area Network) and bluetooth technology; (iv) Microphones and embedded loudspeakers; (v) Keyboard (some models the keyboard are detachable); (vi) Batteries. The Tablet PC’s peripherals allow users interact in different ways, using: hands and keyboard to typewriting; fingers and touchpad to move the mouse pointer; fingers and a virtual keyboard on the touchscreen to typewriting; fingers and touchscreen to trigger an action in a user interface element; hand holding a pen, a pen-sensitive screen and a virtual keyboard to typewriting; hand holding a pen and a pen-sensitive screen to handwriting; voice and recognition software to trigger actions; and other combinations.

Nowadays, computers, tablets, Tablets PC, smartphones, video game console are connected through the Internet. The Internet is a global system of interconnected computer network with standards protocols that allow different technologies to change data. The mainly Web technology, the HTML, evolved and the improvement defined to the last version, the HTML5, are related with support multimedia, keep it easily readable by humans and consistently understood by computers and devices (Berjon et al., 2012). HTML5 adds the new <video>, <audio>, and <canvas> tag elements, as well as the integration of Scalable Vector Graphics (SVG, a vector image format for two-dimensional graphics based on eXtended Markup Language - XML) content and MathML (Mathematical Markup Language is a XML based-format to describing mathematical notations) to integrate mathematical formulae into Web pages. These features are designed to make it easy to include and handle multimedia and graphical content on the web without having proprietary plugins and APIs (Application Programming Interface) installed.

But Web sites and Web applications are still designed to be used with keyboard, mouse (or trackpad in laptop computers) and a high resolution medium sized monitor. One kind of Web application is e-learning (acronym of electronic learning) environments, systems used to support teaching and learning activities that must have good usability and accessibility to support diversity of users and contexts. The e-learning environments had several technological changes in the last two decades, either in changing their user interface, either in the number of available tools, shaped by technological advances and the user’s needs. But the actual version of e-learning environments does not receive input data from many modes to control the application; they only consider trigger actions by keyboard or mouse.

e-Learning environments are designed taking account the user will interact using keyboard, mouse, and high resolution medium-sized screen. But nowadays the users can interact touching a screen with their fingers, with their voices, using a pen, doing gesture with their hands or with body movements. Shneiderman (2002, pg. 21) describe that “The new computing technologies would include wall-sized displays, palmtop appliances, and tiny jewel-like fingertips computers that change your sensory experiences and ways of thinking.” Kugler (2008), supported by reports from Gartner Group, describes one of the big challenges for the Information Technology and Communication (ICT) field in the next 25 years are non-tactile and natural interfaces, and automatic translation of speech. This challenge, coupled with the tendency to change the mouse