



uOttawa

L'Université canadienne
Canada's university

**FACULTÉ DES ÉTUDES SUPÉRIEURES
ET POSTDOCTORALES**



**FACULTY OF GRADUATE AND
POSTDOCTORAL STUDIES**

Ali Karime

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.A.Sc. (Electrical Engineering)

GRADE / DEGREE

School of Information Technology and Engineering

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

The Magic System: An Ambient Edutainment System for the Young Children

TITRE DE LA THÈSE / TITLE OF THESIS

Dr. A.El Saddik

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

W. Gueaieb

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

R. Yul

A. Boukerchel

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

**THE MAGIC SYSTEM: AN AMBIENT EDUTAINMENT
SYSTEM FOR THE YOUNG CHILDREN**

BY

ALI KARIME

A THESIS SUBMITTED TO THE
FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF MASTER IN ELECTRICAL ENGINEERING

OTTAWA-CARLETON INSTITUTE FOR ELECTRICAL AND
COMPUTER ENGINEERING
SCHOOL OF INFORMATION TECHNOLOGY AND ENGINEERING
UNIVERSITY OF OTTAWA

© ALI KARIME, OTTAWA, CANADA 2009



Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-79710-5
Our file Notre référence
ISBN: 978-0-494-79710-5

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

Abstract

This thesis presents an ambient edutainment system that targets the young children and helps them gain knowledge while having fun. The main idea of the system is that children will learn about new objects by either spelling the names of these objects or by tapping over them, and get in response a set of multimedia output related to these objects. For this purpose, we have implemented two approaches. The first approach is based on a voice recognition engine that detects children's speech, maps their unclear words to an appropriate pre-defined match, and provides the system with an accurate term in order to retrieve the appropriate media. Now the second approach is based on the concept of tangible computing. For this purpose, we designed a tangible user interface called "Magic Stick" that reads the tag identifications attached to entities and objects, and then maps them to their pre-defined terms. We incorporated three games that allow children of different ages to benefit from the system's functionalities and encourage them to interact with it.

Acknowledgements

First and foremost, I would like to convey my sincere gratitude and indebtedness to my supervisors, Dr. Abdulmoteleb El Saddik and Dr Wail Gueaieb, for their valuable assistance, comments, and support. A special thanks to Anwar for his long hours of discussions and the great ideas we shared together all over this thesis.

I would like to thank my wife Ghazala, for her inspiration, understanding and support throughout my graduate studies. I also thank my father Mohamad, my Mother Jamal, and my brother Mouemen for their persistent and valuable support throughout my studies. I would like also to express my deep gratitude to all of the MCRLAB mates for their comments and discussions that contributed always to the better of this thesis. Finally, I would like to dedicate this thesis to my daughter Meriam.

Table of Contents

ABSTRACT.....	II
ACKNOWLEDGEMENTS	III
TABLE OF CONTENTS	IV
LIST OF TABLES	VII
LIST OF FIGURES	VIII
LIST OF ABBREVIATIONS	X
CHAPTER 1 INTRODUCTION.....	1
1.1 BACKGROUND AND MOTIVATION	1
1.2 EXISTING PROBLEMS.....	3
1.3 OBJECTIVE AND CONTRIBUTION.....	4
1.4 PUBLICATIONS RESULTING FROM THIS RESEARCH.....	4
1.5 THESIS ORGANIZATION	5
CHAPTER 2 LITERATURE BACKGROUND AND RELATED WORK.....	7
2.1 BACKGROUND.....	7
2.2 EDUTAINMENT THROUGH VIRTUAL REALITY	8
2.2.1 <i>NICE</i>	9
2.2.2 <i>Holodeck</i>	11
2.2.3 <i>Magic Story Cube</i>	12
2.3 EDUTAINMENT THROUGH COMPUTER SOFTWARE	13
2.3.1 <i>FaTe2</i>	14
2.3.2 <i>Kidpad</i>	16
2.4 EDUTAINMENT THROUGH TANGIBLE INTERFACE DEVICE.....	17
2.4.1 <i>I/O Brush</i>	18
2.4.2 <i>Jabberstamp</i>	20

2.4.3 <i>Tangible Flags</i>	21
2.4.4 <i>Jabberwocky</i>	23
2.4.5 <i>E-du box</i>	24
2.5 COMMERCIALIZED EDUTAINMENT PRODUCTS	26
2.5.1 <i>Your Baby Can Read</i>	26
2.5.2 <i>Baby's Learning Laptop</i>	28
2.5.3 <i>LeapSter2</i>	29
CHAPTER 3 PROPOSED APPROACH	31
3.1 REQUIREMENTS OF THE PROPOSED AMBIENT EDUTAINMENT SYSTEM	31
3.2 SYSTEM OVERVIEW	32
3.3 OVERALL SYSTEM ARCHITECTURE	34
3.4 INTERACTION SUPPORT	41
3.4.1 <i>Interaction among Children</i>	41
3.4.2 <i>Children-System Interaction</i>	41
3.5 MULTIMODAL INPUT SELECTION	42
3.5.1 <i>Voice-based Technique</i>	42
3.5.2 <i>Magic Stick Technique</i>	46
3.6 MULTIMODAL OUTPUT SELECTION	51
3.7 SYSTEM'S SCENARIOS	51
3.7.1 <i>Book Game</i>	51
3.7.2 <i>Learning Alphabet Game</i>	53
3.7.3 <i>Object Identification Game</i>	56
3.8 SUMMARY	57
CHAPTER 4 IMPLEMENTATION	58
4.1 COMPONENT AND INTERACTION DIAGRAM	58
4.2 DEVICES AND APIs	61
4.2 GRAPHICAL USER INTERFACE	66
CHAPTER 5 EVALUATION	69
5.1 THE BOOK GAME EVALUATION AMONG PRESCHOOLERS	69

5.1.1 <i>Voice-based Interaction technique evaluation</i>	70
5.1.2 <i>Magic Stick Interaction technique evaluation among very young children ..</i>	74
5.2 THE ALPHABET LEARNING AND OBJECT IDENTIFICATION GAME’S EVALUATION ..	80
5.2.1 <i>The Alphabet Learning Scenario</i>	80
5.2.2 <i>The Object Identification Scenario</i>	88
5.3 SUMMARY	91
CHAPTER 6 CONCLUSION AND FUTURE WORK	92
BIBLIOGRAPHY	94

List of Tables

Table 2.1 Edutainment through VE's work summary in terms of education goals, play space, interface type, and age range.....	9
Table 2.2 Edutainment through VE's work summary in terms of the type of input, output, and setting requirements	10
Table 2.3 Edutainment through computer software's work summary in terms of education goals, play space, interface type, and age range.....	14
Table 2.4 Edutainment through computer software's work summary in terms of the type of input, output, and setting requirements	15
Table 2.5 Edutainment through TUI's work summary in terms of education goals, play space, interface type, and age range.....	18
Table 2.6 Edutainment through TUI's work summary in terms of the type of input, output, and setting requirements	19
Table 2.7 Commercialized product characteristics' summary in terms of education goals, play space, interface type, and age range	27
Table 2.8 Commercialized product characteristics' summary in terms of input, output, and setting requirements	28
Table 3.1 A Sample Mapping table	36
Table 5.1 Results obtained from children's utterance of 75 words	72
Table 5.2 Results observed after the 5 sessions.....	77
Table 5.3 Average time needed for image display of the word uttered.....	79

List of Figures

Figure 2.1 NICE networking connections [1].....	11
Figure 2.2 A user using the cube [3].....	13
Figure 2.3 Kidpad with all local tools shown [5].....	17
Figure 2.4 The I/O brush [6].....	20
Figure 2.5 High level e-du box components [10]	25
Figure 2.6 Baby’s learning laptop [12]	29
Figure 2.7 LeapSter 2 from LeapFrog [13].....	30
Figure 3.1 A high level block diagram of the proposed system	33
Figure 3.2 The overall system’s architecture	35
Figure 3.3 Word Matching Service work flow	37
Figure 3.4 The image filtering algorithm’s work flow	39
Figure 3.5 The structure of the RFID tag mapping.....	40
Figure 3.6 The voice-based interaction flowchart	44
Figure 3.7 The Magic Stick Circuit Architecture [34].....	48
Figure 3.8 The Magic Stick interaction flowchart.....	50
Figure 3.9 The GUI of the Book game	53
Figure 3.10 The RFID tagged keyboard	54
Figure 3.11 The GUI of the Alphabet Learning Game.....	55
Figure 3.12 The GUI of the Object Identification Game.....	57
Figure 4.1 The component diagram of the Magic System.....	59
Figure 4.2 The interaction diagram of the system with the Magic Stick technique	59
Figure 4.3 The interaction diagram of the system with the voice-based technique.....	60
Figure 4.4 The Magic Stick	62
Figure 4.5 The Component diagram of the Magic Stick	62
Figure 4.6 A Google Translate API send request/receive response sample syntax.....	63
Figure 4.7 A Microsoft Agent API send request/receive response sample syntax	64
Figure 4.8 A WordNet API algorithm for requesting words synonyms	65

Figure 4.9 The GUIs associated with the system.....	67
Figure 4.10 The GUI of the ID Mapping table	68
Figure 5.1 The success rate of each of the three trials	74
Figure 5.2 Mary tapping with the Magic Stick on a duck’s picture	76
Figure 5.3 The rate of success of the 4 tapping attempts of the children	78
Figure 5.4 The alphabet tagged board with a pair of Magic Sticks	81
Figure 5.5 A teacher explaining the concept of the Magic Stick.....	82
Figure 5.6 A girl helped her classmate by pointing at the letter B when writing the word Bear	84
Figure 5.7 The output of the word Pizza.....	85
Figure 5.9 The left bottom apple image has caused an atmosphere of laughter among children	86
Figure 5.10 A child asking the teacher if he can write “Pumpkin”	87
Figure 5.11 Success rate of each of the three Attempts.....	88

List of Abbreviations

2D: Two dimensional
3D: Three dimensional
API: Application Programming Interface
AR: Augmented Reality
CCD: Charge-coupled Device
DVD: Digital Video Disc
LED: Light-emitting Diode
GUI: Graphical User Interface
ID: Identification
HCI: Human Computer Interaction
HMD: Head Mounted Display
RFID: Radio Frequency Identification
RGB: Red, Green, and Blue
TCI: Traditional Computer Interface
TUI: Tangible User Interface
TV: Television
UI: User Interface
USB: Universal Serial Bus
VA: Virtual Actor
VE: Virtual Environment
VR: Virtual Reality
XML: Extensible Markup Language

Chapter 1

Introduction

1.1 Background and Motivation

Edutainment, a combination of the terms “Education” and “Entertainment”, has been recently receiving lots of attention and has become a widely emerging research topic. Referring to Druin and Solomon, edutainment can be defined as “a place where children could enjoy what they learn with a combination of many mediums (sound, animation, video, text and images) by simply using a computer mouse to point and click on a particular picture, word, or button; and stories as well as information that will come alive on a computer screen” [21].

Research has revealed the importance of edutainment in improving the cognitive skills of children and stimulating their creativity [20]. Since the early 1990s, interest has risen in developing edutainment systems for children, most specifically applications that possess the temptation of electronic games while accomplishing educational objects [19]. Besides, research has also pointed out that education multimedia systems promise to make learning more enjoyable, convenient and thus more beneficial and effective due to the great potential it provides [22, 23]. Upon these discoveries, there has been a

considerably growing demand to develop edutainment systems associated with multimedia technologies, and that support education and entertainment seamlessly and pervasively [24, 25, 26, 27].

Most of the existing works in the literature approached the topic of edutainment by designing systems and tools for interactive storytelling [28,7]. Other works have developed computer applications for painting and learning games [6]. These systems have encouraged collaboration with other children by various means. Those means were mostly focused on sharing ideas and thoughts, and stimulating discussions among the children [4,7]. However, the usage of these have either required literacy or some basic computer skills, the fact that have made these systems mostly suitable for kindergarteners and older children, and as result have deprived younger children and toddlers of taking advantages of such tools.

Browsing the commercial market, we have found many toys and games that were targeting young infants (birth to 4 years old). Even though these edutainment systems incorporated some learning capabilities, such as teaching new objects, numbers, and alphabets, they have limited the teaching benefits because of their static nature. For instance, infants will only learn a limited set of already defined words, or see the same pictures already associated over and over again.

Investigating the existing systems in both the market and the literature, we have been motivated to design a multimedia edutainment system that will fit the needs of both infants and kindergarteners and that can be used all the way with their various stages of development.

1.2 Existing Problems

The problems of the existing systems are various and differ from one system to another depending on the characteristics and the goals they accomplish. These problems can be summarized in the following points:

- **Entertainment Vs Education Imbalance:** Many systems have been inclined more towards entertainment than education. For instance, systems like [28] and [29] have focused on the fun of children and neglected or did not pay much attention to the education side that edutainment should also provide.
- **Functionalities Vs Price Imbalance:** In order to add more functionalities, some systems had to be equipped with expensive components such as [3] and [7] (HMD, Laptop... etc) which have as a result increased the price of the tool and made it unaffordable for many people.
- **Special System Requirements:** Not all systems could be used in homes or at schools since they require special settings in order to be operated (e.g. Cave). Such requirement have deprive people from using those systems and made them only suitable for experimental purposes.
- **Lack of Interactivity:** Most software based edutainment systems require the children to be sitting for a long time in front of the computer. The lack of interactivity between the user and the system, and between the users have urged many researchers to clarify for the people the bad health, mental, and physical effects that using these systems for long hours would cause for their children.

- **Unobtrusiveness:** The need of using cables among the different components of the systems have created a cumbersome environment for the children and limited their play space.

1.3 Objective and Contribution

The objective of this research is to enhance the edutainment environment by developing new interfaces and functionalities that will make the user's experience more entertaining, beneficial and most importantly interactive.

The contribution of this thesis can be summarized by the following:

- Analyze the requirements of ambient edutainment systems for the very young children and design a tangible user interface (called Magic stick) based on RFID and Bluetooth technologies that do not require any technical skills for them to interact with the target computer systems.
- Design and development of an edutainment system that uses the Magic Stick as a tangible interaction device and facilitates learning for the young children while having fun and being interactive with the system and other participating peers.
- Incorporate multiple multimedia output modalities such as images, audio, and videos that enhance the user's learning and entertainment experience.

1.4 Publications Resulting From This Research

The following 2 papers have been published and they are directly related to the thesis topic.

1. **Ali Karime**, M. Anwar Hossain, Abdulmotaleb El Saddik and Wail Gueaieb, "A Multimedia-Driven Ambient Edutainment System for the Young Children_", ACM MM 2nd Workshop on Story Representation, Mechanism and Context (SRMC08), Oct. 31, 2008, Vancouver, Canada.
2. **Ali Karime**, M. Anwar Hossain, Wail Gueaieb and Abdulmotaleb El Saddik, "Magic Stick A Tangible Interface for the Edutainment of the Young Children", IEEE International Conference on Multimedia and Expo (ICME2009), June 28- July 3, 2009, NY, USA.

1.5 Thesis Organization

The remainder of the thesis is organized as follows:

Chapter 2 presents an overview background literature of the different works that have been done in the domain of edutainment associated with multimedia technologies. It also discusses, if applicable, the drawbacks, and the suitability of the systems for the children.

Chapter 3 proposes our idea and elaborates in details about the different phases of design. The design phases include the voice-based technique, and the RFID-based tangible user interface technique.

Chapter 4 discusses the implementation of the system. This also explains our choice of the different technologies that were used in developing the system, and the sample graphical user interface.

Chapter 5 presents the different test cases we performed, and focuses on the evaluations of system usability and user studies.

Finally in **chapter 6** we summarize our work, draw a conclusion, and propose what might be a fruitful potential future work.

Chapter 2

Related Work

2.1 Background

There have been many efforts made in the field of edutainment, and aimed at promoting learning through entertainment in easy and simple methods. These methods mostly require very low motor skills from the children, and involve a minimal level of computer skills and literacy. Researchers [1,4,6] have tackled the topic of edutainment for children in three different approaches:

- a- By applying the concepts of Virtual reality Environments (VE), for example immersive learning and augmented reality.
- b- By creating computer software that allow children to create their own narrative flow and story elements in collaborative environments.
- c- By exploiting the paradigm of tangible computing through designing different types of Tangible User Interfaces (TUI) for different educational and entertaining activities such as interactive storytelling, painting...etc.

The proposed method outlined in this thesis follows the approach stated in point c. This is due to two principles: the age limit of the target children in mind and the potential of TUI as edutainment tool for this group of children. Nevertheless in the following, all the three groups of works are described as a related literature, with a focus on the third group of literature. Accordingly, this chapter discusses their basic concepts and functionalities. It also, presents some of the existing edutainment toys that already exist in the commercial market. At the end of this chapter, a comparative study between the different approaches stated in these works and our proposed system is drawn.

2.2 Edutainment through Virtual Reality

Virtual Environments (VEs) have great educational potentials since they allow access to the unreachable or the unrealizable [11]. Researchers [1, 2, 3] have exploited the potentials of VEs in edutainment by offering either immersive learning environment where children can interact with the virtual characters and objects or by incorporating augmented reality in storytelling systems. Table 2.1 and 2.2 present a summary of the main characteristics of the works found in the literature in the domain of edutainment through virtual reality in the light of seven distinct points that can be summarized by the following:

- a- Education Goals: these are the goals that each system aimed to achieve in children's education.
- b- Limitation of children's play space: we basically indicate whether the tool requires that the user maintains a static or dynamic position in order to use the system.
- c- Type of interaction interfaces: these are basically the special devices that might be necessary in order for the user to interact or play with the system.

- d- Age range: it shows the recommended children's age range for each work. These ages were explicitly mentioned in the authors' work descriptions and users who undergo the experiments.
- e- Input modality: it actually states the type of action(s) that the user should make in order to activate or interact with the system.
- f- Output modality: it describes the nature of the output(s) that the user will see when using these systems.
- g- Special setting requirements: it indicates whether any type of special places is required in order to operate the system.

Related Work	Education Goals	Is Children's Play Space Limited?	Type of Interaction Interfaces	Age Range (years)
NICE [1]	Explore objects via virtual interaction, share ideas and thoughts via storytelling	No	Handheld, special glasses	6 -10
Holodeck [2]	Share ideas via storytelling	Not defined	Not defined	6 and up
Tangible Story Cube [3]	Improve cognitive skills via augmented reality storytelling	Yes	HMD	Not defined

Table 2.1 Edutainment through VE's work summary in terms of education goals, play space, interface type, and age range

2.2.1 NICE

NICE [1] is an immersive learning environment for children designed to work in the

Related Work	Input Modality	Output Modality	Special Setting Requirements
NICE [1]	Moving handheld	Audio, animation	CAVE
Holodeck [2]	Speech, Object interaction (via an undefined means)	Animation	CAVE
Tangible Story Cube [3]	Folding cube	Animation	N/A

Table 2.2 Edutainment through VE's work summary in terms of the type of input, output, and setting requirements

collaborative Virtual Environment called the CAVE. Based on the current educational theories of constructionism, narrative, and collaboration within a motivating and engaging context, children will be able to construct, cultivate and tend a virtual garden.

NICE allows children to collaborate either together at the same site, or remotely from other located sites. All the users are directly connected through TCP/IP to the garden server running on the NICE web server. Each child is represented by an avatar in the virtual environment, and once he/she speaks or moves, his/her gestures are tracked and transmitted over the network so that other users can see the movements. Through interaction, children can design stories which will be later on published on the web. Figure 2.3 shows the networking connections of NICE among its users.

In order to interact with NICE and to see the virtual and physical world, special glasses are required. In addition, a handheld device is required for interaction with the objects in the VE. Regarding the graphics incorporated NICE uses a Silicon Graphics Onyx with two infinite reality engines.

Obviously, special laboratory settings are necessary to run the environment. The fact

that the system necessitates many expensive tools has made this system unaffordable by the public. In addition, the authors have mentioned that NICE could not be integrated in a classroom setting where basically most of the edutainment systems should be used, and therefore it is mainly used for experimental purposes. It is also worth noting, that this system would greatly be affected by

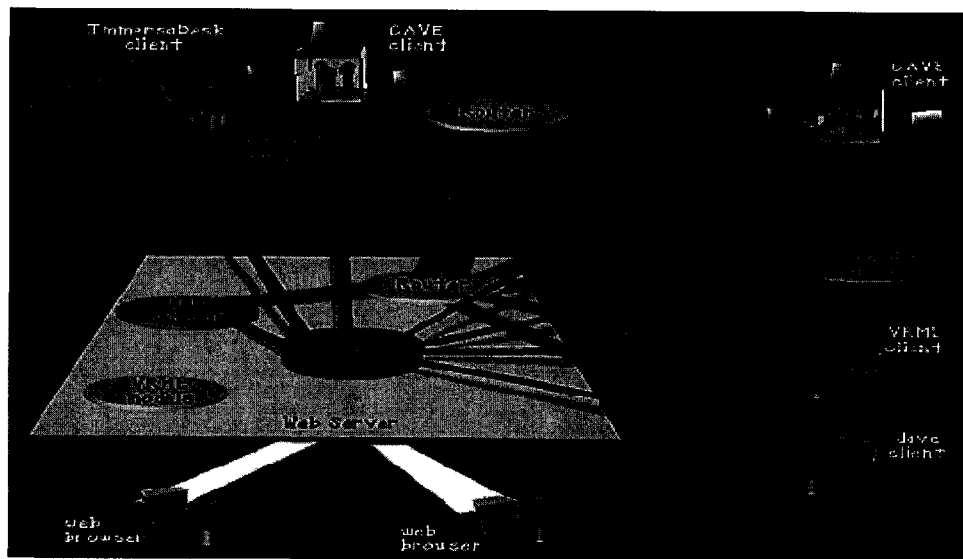


Figure 2.1 NICE networking connections [1]

the number of participants who might cause a bottleneck to the network, and the network delays that might ruin the user's experience while interacting together in the same VE.

2.2.2 Holodeck

In [2], Cavazza has described his ongoing research project and introduced the set of requirements and the fundamental problems behind the implementation of the immersive Storytelling project called the "Holodeck". Similar to NICE [1], Holodeck uses the CAVE technology for VE display. It consists of a number of virtual actors

(VAs) that shares the same environment whom children can interact with. Interaction with the VAs is done through natural language or speech, and the way the user can participate in story generation is in one of the following methods:

- a- Physically by means of object interaction in the virtual environment (VE).
- b- Verbally through on-stage conversation.
- c- Interjectionally through off-stage intervention.

The user's participation might not be permanent depending on his choice. He/she can collaborate in designing the story by participating in the action and then get in or off the stage to watch the scenes of events based on his/her intervention.

Even though the system is still at the very early stage of development, it is not intended to be used in home or classroom settings since it is based on the CAVE. This would make it a very expensive project and mainly intended for experimental purposes. As opposed to our system, we have a cheap, tangible and practical system that can be used at almost all education facilities and homes.

2.2.3 Magic Story Cube

Adopting the concept of augmented reality (AR), the Magic Story Cube [3] aims to provide storytelling for children in a similar way the regular books do but with the association of 3D graphics and audios (speech, music). A cube-shaped specially designed foldable book embedded with multi-sensory devices replaces the regular book that we know. The Magic Story Cube is composed of different transition states that can be reached by folding the cube in a predefined manner. Each state contains some part of the story (drawings). When folding the cube into a specific state, some appropriate segments of audio and 3D animation are played in predefined manner allowing for a realistic interaction feeling with the story. The transition state is Omni-directional, thus the user cannot go back and forth in order to maintain a continuous narrative progression which is a main characteristic in storytelling. Therefore, a transition of one

state to another can be invoked only after the contents (graphics and audios) of the current state have been played.

To see the 3D graphics and scenes, the user has to wear a head mounted display (HMD) with a camera attached in the front. Figure 2.4 shows a girl using the magic Story Cube interface. While the idea of the cube can be seen as innovative, it will not always attract the children's attention and make them interact with the story. In other words, the children will always see the same story all over again every time the cube is being used. This shortcoming will eventually create a feeling of boredom among the children because of the static nature of the tool. In our system, this drawback has been removed by displaying a set of randomly and dynamically updated pictures every time the user explores a new object. In addition, as can be seen from figure 2.4, the HMD on her head limits her movement because of the cable attached to it. It might also cause a feeling of pressure on her head if used for long time.

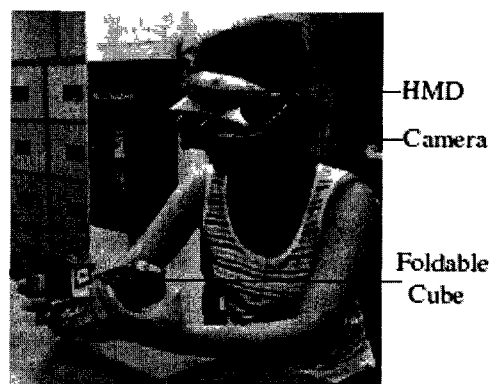


Figure 2.2 A user using the cube [3]

2.3 Edutainment through Computer Software

With the prevalence of computer desktops and laptops almost among all households and school settings, educational and entertaining software have been a key approach for

many researchers. Many types of software have been designed for edutainment and most of them share common characteristics. Some of these softwares exist on almost all Windows Desktop, such as Microsoft's Paint software. Others can be accessed through the web like Funbrain Arcade [16], and the Kidz page [17]. Table 2.3 and 2.4 present a summary of the main characteristics of the works found in the literature in the domain of edutainment through computer software in the light of the seven distinct points mentioned in Section 2.2.

Related Work	Education Goals	Is Children's Play Space Limited?	Type of Interaction Interfaces	Age Range (years)
Fate2 [4]	Stimulate critical thinking via games, share ideas via creating stories, and socialize children via chatting	Yes	Mouse, Keyboard	8-11
Kidpad [5]	Share ideas through collaboration in drawing and storytelling	Yes	Mouse, Keyboard	6-11

Table 2.3 Edutainment through computer software's work summary in terms of education goals, play space, interface type, and age range

2.3.1 FaTe2

FaTe2 [4] is an edutainment environment for young children whose ages range mainly between 8 and 11 years. Based on the advent of hypertext and multimedia, it presents a web based, multi-user, multidimensional hyperspace where children can play, communicate, explore, and build their own stories in a collaborative environment.

Related Work	Input Modality	Output Modality	Special Setting Requirements
Fate2 [4]	Mouse clicking, Keyboard Typing	Audio, animation, Graphics	N/A
Kidpad [5]	Mouse clicking, keyboard typing	Graphics, text	N/A

Table 2.4 Edutainment through computer software's work summary in terms of the type of input, output, and setting requirements

A small group of children (maximum 8), working in the same location or remotely indifferent locations can personalize the scene elements and collaboratively create their own narrative flows. This is done by generating a multidimensional (2 D or 3 D) hyperstory from the linear multimedia stories, which are already built in the system, using a simple authoring tool that supports the creation of multimedia scenes. The system is also empowered with a chat tool that allows sending messages amongst the group.

The system also comprises a set of games that aim at enhancing the motor skills of children and stimulating creative problem solving skills. For instance, linguistic games were incorporated in order to foster the writing skills of children. Other problem solving games, such as “puzzle” game, encourage children to think creatively.

FaTe2 is suitable for elementary school students as the age range suggests. However, the functionalities of the system seem complex to be comprehended by the children. The designers had to associate a virtual human avatar in order to guide the children on how to use the system. That guide was a highly trained teacher who can communicate with the users through the chat area.

2.3.2 Kidpad

Another work that shares common functionalities with [4] is Kidpad [5]. Kidpad is a collaborative storytelling tool for children that supports drawing, typing, and scenes hyperlinking capabilities in a 2D zoomable space. One of its main characteristics is collaboration through accepting input by more than one mouse at a time. This would mean that every child in a group, can paint, zoom or pan simultaneously using his own mouse and in result see his actions, as well as the actions of other children, on the same screen.

In order to be able to distinguish each user's actions, Kidpad uses local tools. Local tools are actually any objects that are graphically shown in the system and that act as cursor. Each tool represents a specific interaction mode, for instance a red drawing crayon can only draw in red at its current location. Figure 2.3 show all the local tools supported by Kidpad. Each user's action is identified through one of these cursors. For example, one child might be drawing using the black crayon, while at the same time another child is zooming or panning. Using Kidpad, children can also draw scenes of a story and link them to other scenes. For instance, users can draw a scene of a person and link this scene to another scene in his head to show what he might be thinking of.

Sharing ideas and stories might be greatly significant for improving the children's mental developments. That's basically what Kidpad mostly focused about. However, we think that including some mechanisms for improving linguistic faculties of the children is also of a great importance. In addition, the main drawback of all software edutainment systems in general is the necessity to be always seated in front of the computer. We are always concerned that sitting in front of computer for extended hours will have dangerous mental, physical and social consequences for children as research has revealed [14, 15]. Our Magic Stick overcomes this important issue since it requires

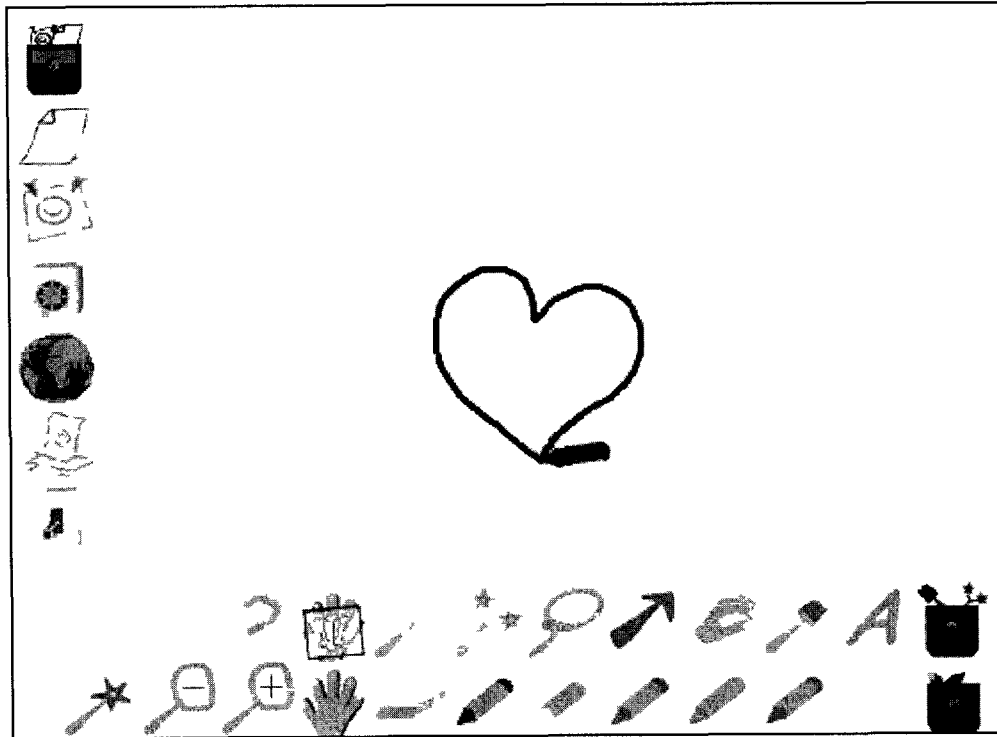


Figure 2.3 Kidpad with all local tools shown [5]

children to interact with the system in different positions (e.g. standing, walking...etc).

2.4 Edutainment through tangible interface device

Since most of the edutainment research addresses young children, it was very natural that designers look for different types of interfaces, other than the traditional ones (e.g. mice, keyboards), in order for the children to have an easy interaction with the systems. Tangible computing promises to make interaction much easier since it allows interaction with the digital information through physical means [18]. Table 2.5 and 2.6 present a summary of the main characteristics of the works found in the literature in the domain of edutainment through TUIs in the light of the seven distinct points mentioned in Section 2.2.

Related Work	Education Goals	Is Children's Play Space Limited?	Type of Interaction Interfaces	Age Range (years)
I/O Brush [6]	Share ideas and thoughts through drawing,	Yes	Tangible Brush	4 and up
Jabberstamp [7]	Share ideas and thoughts via drawing, storytelling	Yes	Rubber stamp, Trumpet, and Microphone	4-8
Tangible Flags [8]	Explore new entities/objects, share information	No	Wireless tablet Computer, Digital pen	6 and up
Jabberwocky [9]	Grammar, vocabulary, spelling ,share ideas via storytelling	Yes	Tangible digital-ink pen	5 and up
E-du Box [10]	Vocabulary, spelling	Yes	Tangible pen	5-6

Table 2.5 Edutainment through TUI's work summary in terms of education goals, play space, interface type, and age range

2.4.1 I/O Brush

Based on creating visual arts to improve the development skills of children, the I/O brush [6] is a drawing tool that picks up different elements of objects and allows children to use these elements to paint their own drawings. Designed in the shape of a regular brush, I/O brush is a TUI, which once brushed over an object, extracts its color, texture, and movement pattern, the elements that will be used later by children for painting.

Related Work	Input Modality	Output Modality	Special Setting Requirements
I/O Brush [6]	Brushing over an object	Still images	N/A
Jabberstamp [7]	Speech through a rubber stamp, pen movement	Audio	N/A
Tangible Flags [8]	Drawing, text , speech	Audio, Text, still images	Outdoor environments
Jabberwocky [9]	Text	Audio, text	N/A
E-du Box [10]	Tangible pen movement	Audio, Still images	N/A

Table 2.6 Edutainment through TUI's work summary in terms of the type of input, output, and setting requirements

The I/O Brush is composed of two main components: the tangible brush and the canvas where the colors of the objects will appear. The brush is equipped with some bendable touch sensors placed at the brush tip, into which 150 optical fibers are woven, to allow for object detection. A CCD camera associated with light bulbs is also mounted on the tip of the brush. Once the brush is swept over the surface of an object, the touch sensors send the activation signals to the bulbs in order to provide an additional light to the camera. Meanwhile, a snapshot of the object is taken by the camera and the frames are sent to the program. Once the ink has been picked up, the fiber optics lights up indicating the ink detection. Figure 2.1 shows the I/O brush's design.

The I/O Brush operates under three modes for picking up the ink: color, texture, and movements. For each mode, a specific process in terms the frames is followed. For example, in case of the texture mode, a snapshot of the brushed surface is captured, which consists of one frame. In the case of color mode, the RGB values of all the pixels in the captured frames are computed and only the most dominating RGB value is

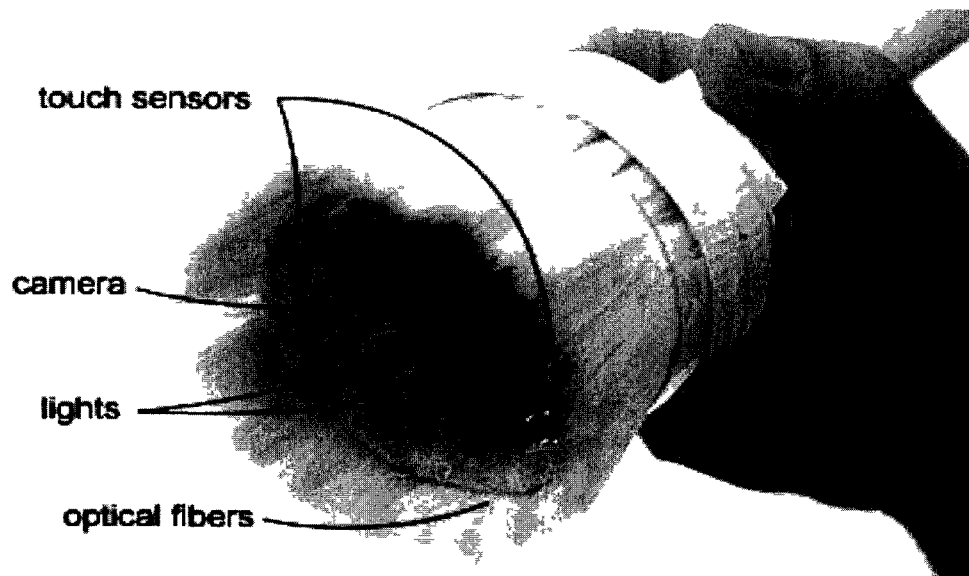


Figure 2.4 The I/O brush [6]

returned so that one color is detected. In the third case, another process occurs that allow children to draw with the moving ink depending on the pattern of movement.

Considering the functionalities of the I/O brush, the tool greatly fits the need of entertainment rather than edutainment since no solid educating evidence was found in the results except that it could be used to teach children about different colors. In addition, the brush was connected to the canvas through a cable that might restrict the movement of the children and limit their play space, which is a drawback we have avoided in our system by using wireless communication among the components.

2.4.2 Jabberstamp

Based on the same concept of visual arts as I/O brush [6], but with the addition of storytelling techniques, Jabberstamp [7] combines the possibility of painting and narrating stories told by children. Designed in the shape of a rectangular tablet with a regular sheet of paper on the top, Jabberstamp gives children the illusion that their

voices are coming from the paper. The system consists of a Wacom tablet, two modified Wacom pen, pre-printed paper templates, speakers and microphones. The components are connected to a computer where a custom software is running.

The tool can operate in two modes: “single page” and “book” mode. In the single page mode, children can paint their own drawings on one sheet of paper placed on the special tablet and record their stories about their drawings by pressing a self-inking rubber stamp onto the page. In order to hear the recordings, they place a special trumpet on the marks of the stamp. Hence, children will be allowed to recall the stories they have narrated about paintings they have drawn in the past depending on the context at that time. The context might be the mood of the child at the time he told the story or the thoughts he had at that time about his paintings. In the book mode, multiple sheets can be used and each drawing sheet is assigned a specific number to allow for playbacks whenever a child places an old drawing. The sounds and audio recordings are saved on a computer, so that at any point in time it is possible to retrieve the pages’ appropriate audios.

Not only is Jabberstamp considered as an edutainment tool, but also it can serve as a communication tool since it conveys to teachers and adults meanings depending on the page’s illustrations that children mark. Jabberstamp can be used by children with minimal literacy; however, a certain understanding level of objects interaction is necessary for proper usage. Therefore, only children of the age of 4 and up might be able to use it. In addition, the form of the tools did not automatically imply their use to the children as the author have mentioned. In our system, such confusion is eliminated since children will be dealing with a Magic Stick that is easy to use, and that doesn’t require much demonstration before usage.

2.4.3 Tangible Flags

In [8], Chipman presented his Tangible Flags that are used for collaborative learning in

outdoor trips. The system provides children with the opportunity to explore the outdoor entities, collaborate by sharing their ideas and knowledge, and construct digital knowledge artifacts while they are in outdoor field trips.

The system consists of a set of RFID tags embedded in small flags, each assigned with a specific color. A RFID reader attached to a wireless handheld tablet computer is used to access the digital information saved on a remote server about each of the tangible flags.

Children will be asked to place these Flags in the outdoor environment they are in; for instance, they will be asked to find some yellow flowers or trees in the garden and place these Tangible Flags beside them after they save their ideas and thoughts in each of the flags. Scanning a tag will allow the user to gain access for the digital information mapped with that specific tag. By accessing the information on a specific tag, a group of children will be able to jot down their observations and discussions about the environment around that Tangible Flag. They can draw pictures, record audios or videos, and scribble their inquiries. When another group of children comes after to explore that same area, they can use their wireless computers to scan the Tangible Flag already accessed by the first group. Then, they will have the opportunity to read the information or paintings saved by the first group. They can also add their own comments and thoughts about that same environment, answer any queries, hear or add any new recordings. However, they will not be able to modify or delete any piece of information already saved by the first group, so that every consecutive group will be aware of the different opinions and ponderings of all the previous visiting groups. The process start over again every time a new group scans that specific flag, which will create a collaborative educational environment among children.

The fact that Tangible Flags system imposes the requirement of carrying a tablet computer might be exhausting for the children. This fact has encouraged us to eliminate

this inconvenience by carefully designing a light weight stick that could be used by children of all ages for extended hours.

2.4.4 Jabberwocky

Jabberwocky [9] is a system that concentrates on the importance of improving children's writing and grammar skills by recording stories told by them and correcting their mistakes. It provides a means to record and share written stories with other children in a humorous and entertaining manner. The basic idea behind Jabberwocky is that children can improve their writing skills by writing their thoughts stories using a digital pen and paper. Then the text will be checked for writing errors, and depending on a certain algorithm their mistakes will be either "Jabberwockied" or modified.

The Jabberwocky's interface consists of a small camera mounted on the tip of a digital pen. The camera is meant to capture the stroke information when the user is moving the pen over a digital paper, and consequently provides a replica of the ink writing that the user has created. The system operates in two modes: a free unconstrained version and a constrained version. In the former mode, the system will replace any non-word generated by the recognizer, such as numbers or word that doesn't exist in the dictionary, by the word Jabberwocky. Now in the latter mode, the system might or might not replace the non-words by jabberwocky depending on an acceptable error rate set by the teacher. It might as well perform a scan in the lexicon and replace a miswritten word by a closest match. If the error rate of the child is higher than the customary rate, the system won't even write the text of the story and produce instead a "failure to read message" so that children won't feel discouraged because of their high error rate. Finally, the modified story will be spelled by a speech to voice system so that everybody could share other's stories.

The system is mainly used in school settings to help improving the skills of the children of at least 5 years of age. Though Jabberwocky has a high potential for children's education, it suffers from some drawbacks. Firstly, in the first mode, the system will always replace the errors by the word "Jabberwocky" instead of showing children the correct way of spelling these mistakes. For instance, if a child made 10 writing mistakes in a 20 word text, the system will display a text with half the words jabberwockied, which is an undesirable event that might cause confusion among the children. Secondly, in the second mode, the words modified in the text might not necessarily be the words that the child really meant to write since these words are sometimes replaced by the closest match. This would sometimes result in changing the whole meaning of the text.

2.4.5 E-du box

E-du box [10] is an edutainment system associated with a tangible input pen and a tangible, interactive, and animated agent who can move and speak to the users. It is a hardware/software educational platform that addresses mostly children between the age of 5 and 6, especially those with no computer skills. It allows learners to experience content created by their educators in an intuitive and stimulating way by exploiting the benefits of multimedia and tangible interfaces.

The system is composed of the E-du Box device and three other main parts:

- a- A special pen-shaped mouse that can vibrate according to the educator's need.
- b- An interactive agent who can speak to users.
- b- An educator computer.

All devices are connected wirelessly with Bluetooth in order to provide a seamless environment. The E-du box makes use of the ordinary TV to display the output. As soon as the E-du box device is connected to the TV terminals, a graphical user interface in the shape of a traditional notebook is displayed. The notebook contains games and assessments associated with a rich multimedia experiences such as images and spelling.

Different educational interactive exercises are built in the system such as Draw Together, Memory Test Game, and Circle picture. Most of these games or exercises aim at testing and improving the knowledge of the users in a simple way. For instance, in Circle Picture, users, working individually or collaboratively, will be asked to circle a picture that starts with a certain alphabetical letter. Depending on the user's choice, the external agent will congratulate the user in case of a right answer, or ask him to try again if the choice is wrong. These activities are monitored by an external computer used by the teacher or educator who can synchronize with students, and check their progress. One of the interesting features the system provides is that it allows educators to develop their own multimedia interactive exercises or games. Therefore, using the information they have collected during the exercises and the observations regarding their students' progress, teachers can create their own E-du box exercises based on the needs of their students. This is accomplished by using a powerful authoring tool, called the "E-ducreator" that

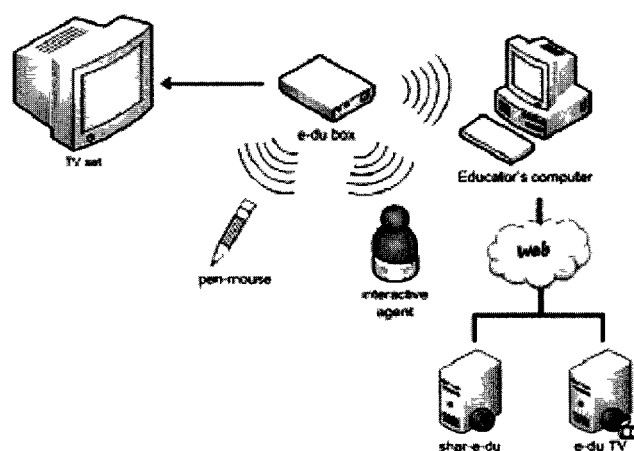


Figure 2.5 High level e-du box components [10]

runs in the educator's computer. Figure 2.2 shows the different elements involved in the e-du box system.

The authors have stated that while testing, children have faced problems using the digital pen as an input device [10]. The most crucial problem was that children were sometimes lifting the pen from the table's surface which has caused a faulty interaction between the user and the computer. Even though the connection between the pen and the e-du box tool was wireless, this has limited their movements since to be able to correctly interact with the system, they had to work on the surface of a table.

2.5 Commercialized edutainment products

There are vast numbers of toys and educational tools for children that have already been commercialized and became very popular almost in all children's homes. These toys are classified depending on children's age ranges and their functionalities are various. The price of these tools also ranges from as low as few dollars to as high as hundreds of dollars depending on the capabilities they provide. Table 2.7 and 2.8 present a summary of the main characteristics of the works found in the commercial market in the light of the seven distinct points mentioned in Section 2.2.

2.5.1 Your Baby Can Read

Targeting very young children, "Your Baby Can Read" [11] is an early language development system that aims to teach the language and phonics for young toddlers by showing a set of words, teaching them how to spell them, and show related images describing the meaning of those words.

Parents will be provided with some set of DVDs that can be seen using a DVD player on a TV screen. One the DVD is played, a word will be shown. At the same instant of the word display, the infant will hear a sound of someone spelling that word. And finally, in order to consolidate the word in the child's memory, a predefined picture that

might describe the meaning of that word is shown, For instance, if the child sees the word “clap”, an image of a monkey clapping with his hands will be displayed.

Related Work	Education Goals	Is Children's Play Space Limited?	Type of Interaction Interfaces	Age Range (years)
Your Baby Can Read [11]	Teach language, phonics, learn about new objects	Yes	N/A	0.25 - 5
Baby's Learning Laptop [12]	Shapes, objects, some words	No	Baby's Learning Laptop toy	0.5 and up
LeapSter2 [13]	Spelling, simple math, share ideas via drawing	No	LeapSter2 toy	4-8

Table 2.7 Commercialized product characteristics' summary in terms of education goals, play space, interface type, and age range

This system lacks the characteristic of interactivity. The system requires children to watch the television screen all the time, the fact that makes us pose a crucial question. Since we cannot force children not to move from one place to another, the first question might be on how to attract children's attention all the time on the screen if there is no interaction device or a tool to interact with the system. Moreover, wouldn't children simply get bored just by watching? We have simply avoided such questions by designing a TUI that will allow a fun way of interaction and that doesn't require children to stand still in one position.

Related Work	Input Modality	Output Modality	Special Setting Requirements
Your Baby Can Read [11]	N/A	Video, Still images, audio	N/A
Baby's Learning Laptop [12]	Button press	Audio	N/A
LeapSter2 [13]	Button press, pen movement	Graphics, Audio	N/A
Our system	Tapping on an object/Tag, Speech	Audio, animation, still images, text.	N/A

Table 2.8 Commercialized product characteristics' summary in terms of input, output, and setting requirements

2.5.2 Baby's Learning Laptop

Since children would like to imitate their parents, “V-tech Electronic Learning” [12] has designed a laptop toy for infants of the age of 6 months and up. Structured in the shape of a multicolored small laptop, Baby's Learning Laptop attracts the infants' attentions with the screen lights produced every time a button is pressed. Figure 2.6 shows the design of the laptop. The toy works in three different modes. The first mode teaches either the shapes of the buttons that the infant presses (e.g. oval, rectangle) or the common name or sounds of the drawings on those buttons. For instance, the word “horn” will be uttered if the button where a horn is drawn is pressed, accompanied by multiple lights flashing on the screen. The second mode gives some kind of description of each drawn object on the screen. Once the child presses a button, only one light of that particular drawing flashes with a small sentence regarding that object. The third mode is meant to give children some fun while providing a set of pre-saved songs, music, and sounds once a button is pressed.

The “Baby’s Learning Laptop” is actually a very nice and entertaining toy for very young children. However, it would be much more beneficial in our opinion if visual representations are shown, since it would greatly influence not only the interaction of the children with the toy, but also the learning experience. It is also worth mentioning that such toys would suits the needs of pre-school children only. For kindergarteners, some more advanced toys are required.



Figure 2.6 Baby’s learning laptop [12]

2.5.3 LeapSter2

LeapSter2 [13] is a portable edutainment system from LeapFrog that aims at promoting learning through games. Figure 2.7 presents the LeapFrog2 portable system design. The games in LeapSter2 are specially designed to improve children’s reading skills by focusing on the phonics skills, consonants, sight words, homophones, and vowels. It also has various games that allow children to explore different shapes, such as triangles and rectangles, and to teach them some basic math operations, such as addition and subtraction, through eliminating or adding objects in the game.

LeapSter2 can be also connected to the computer through a USB, and then they can upload their paintings to an online creativity studio where they can embellish and share their drawings with other children. Parents can also benefit from the online facility of the system by accessing the online LeapFrog Learning Path where they can see the progress of their children, check which activities their children are mostly engaged with, and check the type of questions their children are exploring.



Figure 2.7 LeapSter 2 from LeapFrog [13]

Chapter 3

Proposed Approach

In this chapter we draw the requirements of our proposed system and discuss the overall architecture with the different modules associated with it.

3.1 Requirements of the Proposed Ambient Edutainment System

We have set the requirements of our proposed ambient edutainment system based on the existing problems discussed in section 1.2 of chapter 1 and our objectives mentioned in section 1.3 of the same chapter. Research has been conducted in order to identify the most significant characteristics that the system should possess in order to contribute to the education and entertainment of children and improve their various development skills. Finally the requirements that we came up with are the following:

1. **The system should be interactive:** Our system should involve children in a dynamic play environment and engage them into discussions and solving games with other peers. All these can be supported if the edutainment system provides mechanisms so that the children do not have to just sit in front of a computer to learn things or play games.

2. **A Multimodal Interaction Interface should be available:** Keeping in mind that the system is targeting children of different ages, we need to consider their different motor skills that would allow them to interact easily with the system. Research has indicated that the gross motor's control is far better developed than the fine motor control for children at very young ages, which will make it very difficult for them to guide and click a mouse. Children are expected to be able to use a mouse sometimes during the preschool years [33]. For this reason, an interaction interface that might be appropriate for one child might not be for another. Therefore, more than one type of input interaction interface is vital.
3. **A multimodal Output display should be incorporated:** Researchers found out that an increment in young children's performance has been detected when pictures were added to orally presented stories [38]. Moreover, studies pointed out that the use of different media such as, audio, video, text, animation and graphics have a tremendous potential for improving the learning skills [39]. Based on those facts, our system should provide more than one type of multimedia output so that it could contribute to the learning and entertainment of children in an effective way.

3.2 System Overview

The core idea of the proposed system is that whenever a child touches an object or utters the name of an object or entity, some appropriate media representations such as audio, image, video, and texts are presented. Based on this initial idea, we have developed the system that operates based on voice and RFID-enabled inputs and

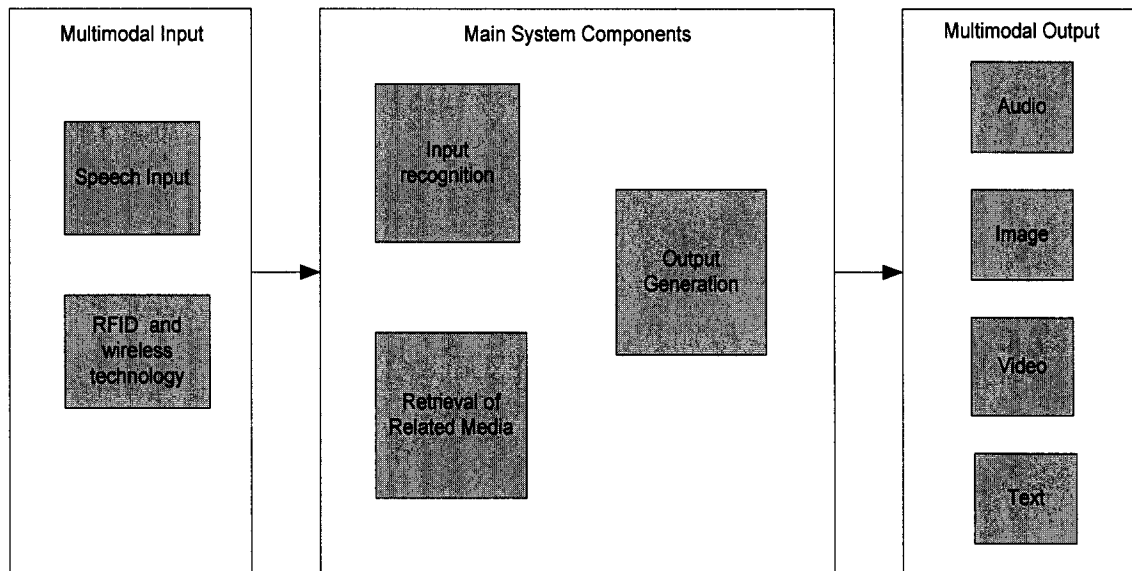


Figure 3.1 A high level block diagram of the proposed system

provides appropriate visual outputs. The RFID-based input is supported by a tangible user interface (TUI) tool, which we developed for easy interaction. To demonstrate the system's functionality, we also developed three specially designed games that are suitable for children. All three games actually share the common characteristics of showing images or videos from the web and/or from a database, uttering the words or letters that the children say, and translate the uttered words into other languages of children's or guardian's choice. The three games are: Object Identification, Alphabet Learning, and a Book game.

The "Book" game targets infants, while the other two are suitable for children of all ages. The system requires no computer skills and very low literacy level to be operated and will use a natural way of human computer interaction (HCI) through the use of two techniques of interaction interfaces: the voice enabled interaction and the RFID-based interaction.

The first technique is based on the regular voice recognition technique associated with some methods we created that help the system identify the ambiguous input words uttered by very young children. The second technique is based on the ease and simplicity of interaction using TUIs. We have created for this purpose a new TUI, which we called the Magic Stick. It uses RFID and Bluetooth technology and helps children interact with the games. Both techniques might be used for the same purposes; however they cannot be used simultaneously. We discuss the two techniques in more details in the subsequent sections. Figure 3.1 presents a high level block diagram of the overall system, which shows that the system supports multiple input modalities that are processed to provide multimodal output in the environment.

The remainder of this chapter is organized as follows: Section 3.3 presents the overall system architecture and discusses its various modules, Section 3.4 states the interaction support of the proposed system, Section 3.5 discusses the two input interaction techniques and their implementations, Section 3.6 illustrate the multimodal output generation techniques. Finally Section 3.7 elaborates the different game scenarios associated with the system.

3.3 Overall System Architecture

We now provide an overview of the overall system architecture. Figure 3.2 presents the architecture of the overall system including the different services needed to realize its functionalities. We briefly explain each of these services in the following:

Voice Recognition Service: One of the two types of input interactions is the voice-based technique. When the children are interacting through speech, the Voice recognition service will detect the speech of the user. It is developed using Microsoft Speech Recognition Engine [46] that detects the input speech as word. The detected word will be sent to the Word Matching Service for further processing.

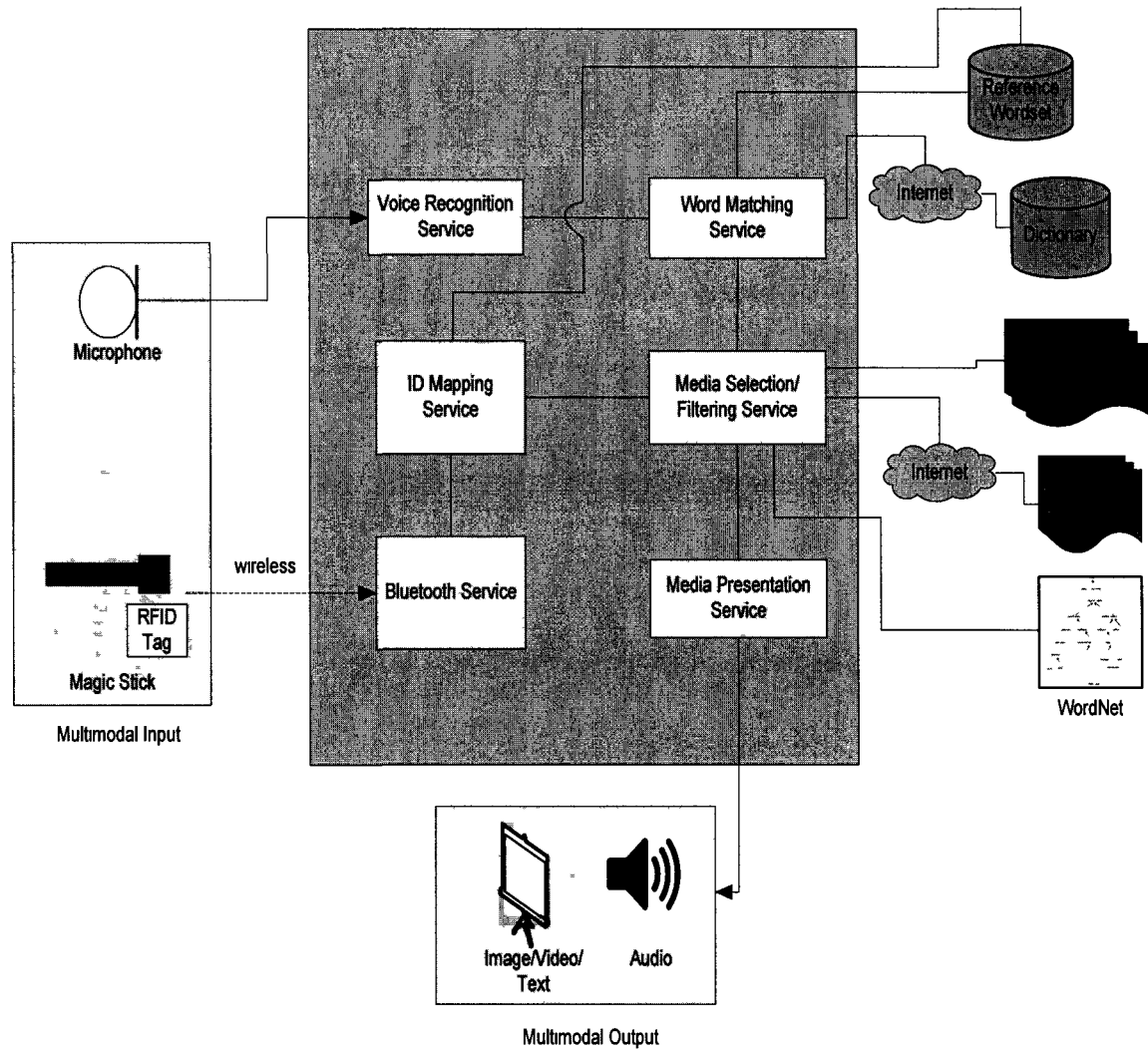


Figure 3.2 The overall system's architecture

Word Matching Service: The Word Matching service maps the input word with an appropriate match by checking the reference wordset database and the dictionary. The reference wordset database consists of a 3 column mapping table (see Table 3.1) that contains information about the words the child can speak (Entity/Person name), the way

the child spells or utters these words (child's spelling), and a flag bit that specify whether a certain word is the name of a person (proper name). The online dictionary allows to determine whether a word that is not entered by guardians in the wordset database exist in the English language or not. Figure 3.3 presents a diagram of the Word Matching Service work flow.

Entity/Person name	Child's spelling	Flag bit (1 for proper, 0 for improper)
Banana	Bana	0
Beeb	Milk	0
Joe	John	1

Table 3.1 A Sample Mapping table

In Figure 3.3, it is clear that upon receiving the detected word from the Voice Recognition Service, the Matching Service first checks if the word exists in the first column (Entity/Person name) of the mapping table (Table 3.1). If it exists, it will pass to the next service (Media Selection) its equivalent mapped word that exists in the second column (child's spelling) and the flag bit as either 1 or 0 that indicates whether the word is a proper name or not, respectively. For instance, if the voice engine detects the word "Bana", the matching module will determine that its equivalent word is "Banana". Now, if the word detected doesn't exist in the first column of the table, then the Matching service will check if that word has a match in the online dictionary. If a match is found, the word will be passed to the next module. If no match is found, the Matching service will notify the Media Selection module that the word doesn't exist, and consequently the system will end up displaying a random cartoon image from the personal image repository.

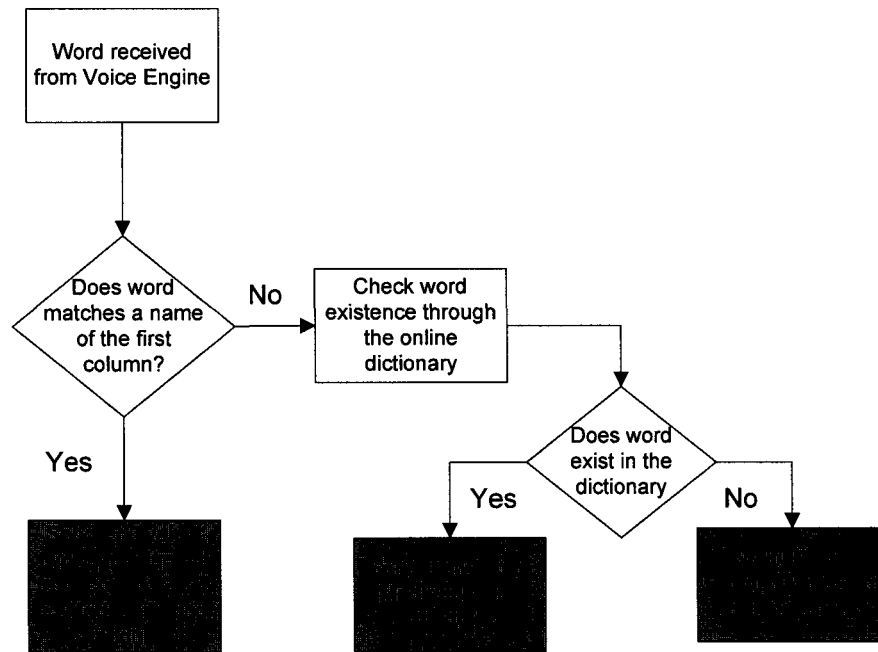


Figure 3.3 Word Matching Service work flow

Media Selection and Filtering Service: The Media Selection Service selects the appropriate picture to be displayed by the system. The selection depends on the type of input received from either the Word Matching service or the ID Mapping one (described later). In case of flag bit 1 input received from the Word Matching Service, appropriate picture related to the name of the person being uttered will be searched and retrieved. Now if it's case of a non-existing word, it will randomly select an image from the personal repository and pass it to the Media Presentation service for display. Finally in the case of an English word, the service will fetch some appropriate images related to that word through a webservice and ensure that no offensive materials are displayed, which is achieved through filtering.

It is obvious that filtering the content is of a big importance especially when dealing with the children. We have adopted Rahman's [41] search result clustering scheme that is based on the semantic lexicon WordNet [42] that groups the English words into a set of synonyms and calculates the semantic relations among the synonym set. We have used the semantic relation calculation method and created the offensive word filtering process and added that in our system. By using the semantic distance calculation method between two given words (A and B for example) we have defined a threshold $\alpha(A, B)$. The threshold gives us the minimum semantic distance between A and B , where we can deduce that they are generated from the same parent. In the Hyponym relation tree this distance is usually less than three. For, precautionary reason, we have considered the value to be four. This threshold simplifies the design of the overall process. A flowchart that represents the filtering procedures is shown in figure 3.4 for clarification.

Each retrieved image on the web repositories (e.g. Yahoo, Flickr) is associated with a set of attributes which are the tags, title and description. The offensive words could reside in any of these attributes. The process starts by first removing from the attributes the words such as *has*, *is*, *he*, which do not have any significance in the filtering process. Removing these words will limit the attributes of the images and will provide us a set of keywords that describe the image. Hence we can formulate our algorithm to process these keywords to examine whether they are offensive or not. In order to accomplish this, we have created a sample of sorted list of known offensive words. The sorted list is helpful when we perform binary searching operations over the list. Now for each refined attributes of the picture, first we binary search them in the offensive word list. If a match has been found we stop processing the keywords and classify the picture as being offensive. Otherwise, we generate synonyms of each keyword using the WordNet library [42]. Afterwards, we calculate the hyponym distance [41] between the synonyms of the keywords and the known sorted offensive word list we have previously defined. We finally compare the search with above-mentioned threshold α so that we guarantee

that the synonyms are not inherited from any of the parents of the known sorted offensive word list. If the synonyms passed the thresholds we conclude that there is less possibility of the image being offensive and therefore it can be selected for display.

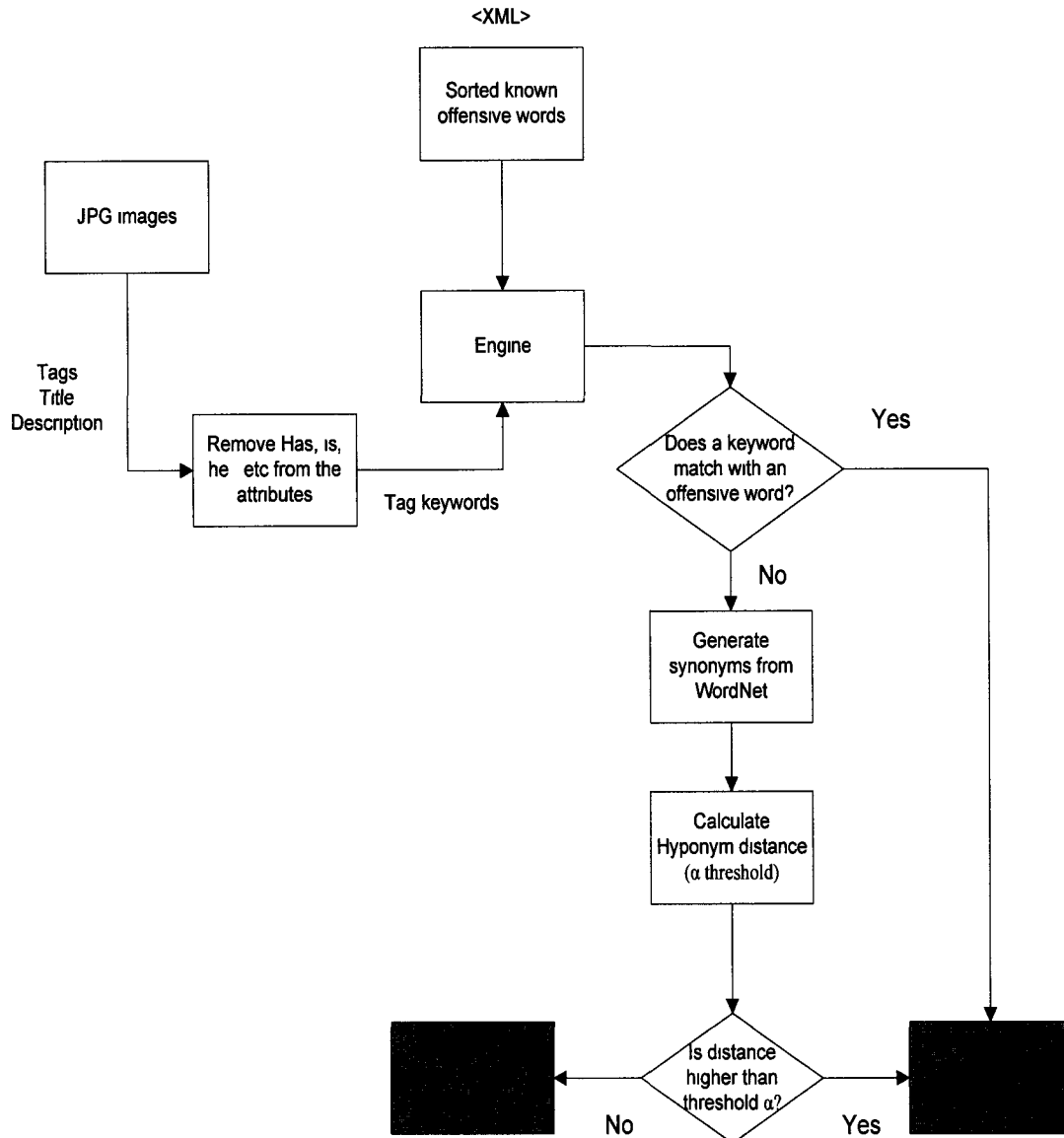


Figure 3.4 The image filtering algorithm's work flow

Bluetooth Service: This service is activated when the system operates with RFID based input. The Bluetooth service consists of the Bluetooth dongle that is responsible of receiving the wireless data sent by the Magic Stick.

ID Mapping Service: The ID Mapping service is responsible for matching the Tag IDs received from the Bluetooth service with the appropriate mappings initially entered by the guardians in the reference wordset database. We have created a 5 column mapping that can be modified depending on the guardian's choice. The XML structure of the table is shown in figure 3.5. The first column is for the tag ID. The second column is the mapping between the tag ID and its associated mapping word (e.g. apple). The third column is used for specifying the type of the objects or entities (e.g. bird, fish, fruit). The fourth column contains a description related to a tag. This description can be provided by the guardians to be uttered by the system. Alternatively, the description can be optionally extracted from the web resources. The fifth column keeps the image/video location associated with a tag. Upon finding the appropriate mapping, the ID mapping service will provide the Media Selection/Filtering service with the mapped words of the second columns so that appropriate images could be fetched.

```
<tags>
  <tag ID="ABCDEF000123"
    <wordMap></ wordMap >
    <type></type>
    <description></description>
    <imageLocation></ imageLocation >
  </tag>
</tags>
```

Figure 3.5 The structure of the RFID tag mapping

Media Presentation Service: After receiving the appropriate media from the Media Selection Service, the Media Presentation service will display the related images, audio, video, text or animation depending on the game scenario and the guardian custom settings.

3.4 Interaction Support

The proposed ambient edutainment system supports two modes of interaction, which are interaction among children and interaction between children and the system. These are described in the following:

3.4.1 Interaction among Children

We have provided interaction support among the users by associating games and quizzes that encourage, depending on the scenario, discussions among them, sharing opinions, and sometimes collaborating on how to solve a certain task. For instance, children can help each other to find different types of birds from a number of doll objects scattered in their classroom using the Magic Stick (the design of magic stick is described in a later section). In another example, children can write together a name of a fruit using the Magic Stick and get in response some multimedia representations related to that fruit from the web.

3.4.2 Children-System Interaction

The proposed system provides mechanisms such that the children do not have to sit in front of the computer to interact with the system unlike the traditional computer interfaces (TCIs), which require children to use mice and keyboards that limit the

interaction of the users. Interaction between the users and the proposed system is achieved by choosing the appropriate means of input interfaces (e.g. speech and Magic stick). This approach does not limit the children's play space and allow them to interact with the system regardless of their positions. They could be standing, walking, leaning, or sitting to provide input to the system and yet get the response using appropriate output modalities.

3.5 Multimodal input Selection

We propose a multimodal input capability that will give the child more options to interact with the system and to allow the guardian to pick what he/she thinks is the most appropriate means of input interaction depending on the child's age. Two techniques of input interaction interfaces are proposed:

- 1- The voice-based technique that will allow the children to interact with the system through natural speech.
- 2- The tangible interface technique that will provide the children with a special physical interface called Magic Stick that allows them to interact with the system remotely in a natural way.

3.5.1 Voice-based Technique

The idea behind the voice-based technique is that whenever a user utters a word, he/she will be provided with an appropriate output related to that word. The fact that young children will be speaking to the system imposed some technical challenges on how to identify the children's ambiguous words that they might say which can be stated as follows:

- a) The speech recognition systems that exist in the market or research community suffer from inaccurate speech detection [45]. They also work in controlled environment to obtain better performance. In addition, they require some training to produce better results.
- b) Unlike adults who can be instructed on how to properly speak with the system, children cannot be forced to speak in a particular way. Therefore it is difficult to provide speech based input facility when dealing with the children. Moreover, children might always use some words that might not have a meaning or exist in the dictionary. Consequently, we needed to adopt a mechanism that will help the system identify the unclear words uttered by children. We discussed this mechanism in Section 3.3 of this chapter.

Interaction with the voice-based technique

The usage of the voice-based technique depends on the clarity of the user's speaking. For children who can clearly speak, the system can be just used anywhere where technology allows. However, if it is the case of very young children who are at an early speaking stage, the presence of the parents or adults who have good knowledge about the children is necessary in order for the system to operate properly. This is due to the requirement of setting up the predefined words and different images or videos to the system as well as the limitation of speech recognition system.

In order to clarify how interaction is done with the voice-based technique, we present the flow diagram in figure 3.6 [35]. Using this technique, the system accepts the words uttered by the children as input. Since children might say more than one word, we had to avoid such confusion by limiting the input to the first word detected. Once the system detects a word, the speech recognition engine will stop working until the expected output is displayed.

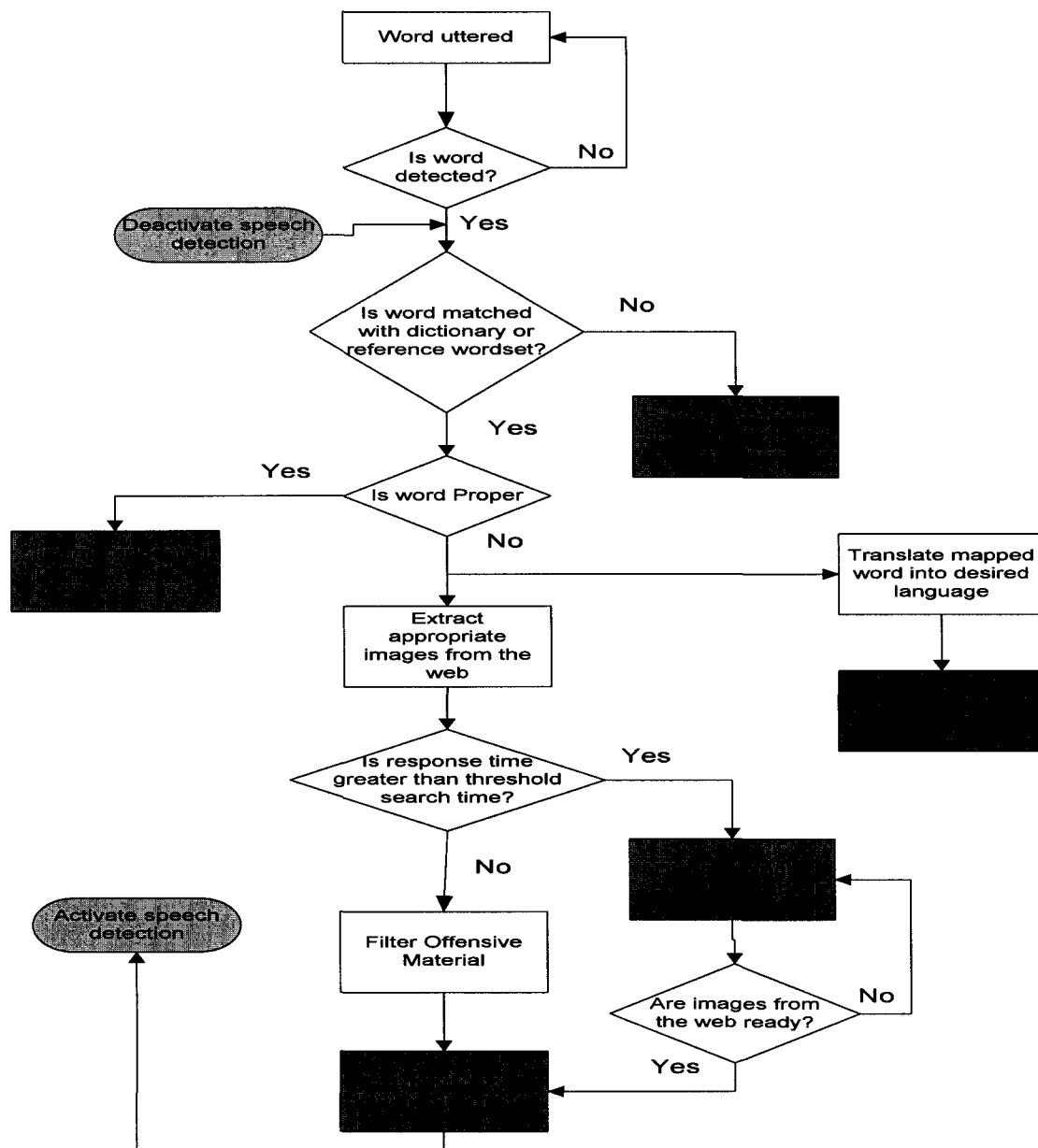


Figure 3.6 The voice-based interaction flowchart

Now, once a word is detected, the system will check if that word matches any of the words in the reference wordset. The reference wordset is basically the database where parents map the words their children can spell, based on their children's accents or their

tones of pronunciation, with the real word. For example, if parents know that their child spells the word “banana” as “bana”, they can map the words in the database so that it becomes understandable to the system. This concept will be explained in more details when discussing the architecture. If no match is found in the wordset, the system will query an online dictionary to check if that word exists in the English language. If the word has a meaning, the system will fetch for appropriate images from the web related to that word, unless this is an objectionable word. If the word doesn’t exist in the dictionary, then in this case it won’t have a meaning and a random cartoon image from the database is displayed so that the child feels motivated to continue the process of speaking.

The possibility of a children uttering of a person (proper name) is valid. The word detected might always be of any type (verb, proper name, adjective). Querying the name of a person through the web might result in thousands of different images for people with the same name. In addition, the child might utter the names of his parents or a member of the family. For this reason, we have given parents the option of saving some pictures in the database of the people whom their children can utter their names. In this manner, if the word detected is proper, an appropriate image from the database could be displayed.

Displaying images from the web could not happen before checking their contents. Filtering is a very crucial step at this level so that no offensive material is displayed. The filtering procedure was discussed in details in Section 3.3. In case of a long process of image fetching or filtering delays, the system will display images from the database and check if the images requested are ready to be displayed. Once images from the web are found and ready, the system will change the existing displays by dynamically displaying the images from the web. At the same time of display, the word will be spelled in the proper manner, written and shown on the screen, and translated to the language of choice.

3.5.2 Magic Stick Technique

To overcome the limitation faced by the speech-based input, we have designed and developed a TUI tool, which we call Magic Stick [34]. The main goal of using the Magic Stick is to allow children to interact with the digital information through a physical interface in a natural way. The core idea behind it is that children will be able to perform actions through tapping over tagged objects. For instance, to learn about new objects/entities in their surrounding environments, children will only have to tap these images and audio pertained to the objects being tapped. Before designing the Magic Stick we had to impose some set of technical requirements that would facilitate its use among children of all ages and simplify their interactions with the system. These requirements could be resumed as follows:

- **The Magic Stick must to be unobtrusive:** In order to achieve the first requirement mentioned in Section 3.1 in this chapter, we had to avoid using components that require a cabled connection among the Magic Stick and the computer. In order to fulfill that, we have used a wireless connection using Bluetooth communication that will send the data to the computer from as far as 100 meters by using a Class I Bluetooth chip that provides such transmitting distance. With a wireless connection, children will enjoy a far better flexibility of movement than in a wired connection.
- **The Magic Stick should assume no technical skills or literacy level:** Based on the fact that the system is targeting users of all ages, especially the children, the Magic Stick should not require any specific technical skills to operate. Contrary to the mouse that requires the user to click and move the cursor, or to the keyboard that needs some literacy level to be used, the Magic Stick only requires the user to hold it and touch entities.

- **The Magic Stick must be light:** Keeping in mind that the Magic Stick will be carried by the children for an indefinite amount of time, it was obvious that the weight has to be carefully considered. We have used in the design of the Magic Stick as simple electronic components as possible in order to reduce the weight of the tool and at the same time eliminate any burden of carrying it for long hours. The electronic components were mounted on a light wooden stick. We covered the stick with a hard carton and placed sticker on top of it to attract children.
- **The Magic Stick should be a safe tool:** A very important aspect is of course the safety of using the Magic Stick. For this reason, we have chosen the technologies that were deemed to be safe by experts.

The Design of the Magic Stick

Our proposed TUI uses RFID technologies for object identification as this technology has already been approved by experts for use in humans since it is considered as a safe and effective tool in a variety of applications [36]. We have placed an RFID reader on the tip of the Magic Stick that will allow reading the digital information from the RFID passive tags which will be placed in the different objects and entities. Using passive tags eliminates the need for a power source and therefore these tags could be placed into any object without any special needs.

For communication among the Magic Stick and the computer, we have used Bluetooth due to its secure and low interfering wireless connection. Researchers agree that due to the low power levels required to operate, Bluetooth-enabled products have not been identified as posing any health risks [37]. A Bluetooth chip that accepts a serial data stream was connected with the RFID reader. On the computer side, a Bluetooth USB

dongle was attached in order to receive the serial data transmitted by the Magic Stick, which is used by our system to provide the required output.

In order to eliminate any specific knowledge on how to turn on and off the tool, we have integrated a proximity sensor that will either turn on or off the Magic Stick depending on a specific distance between the user and the interface. The Magic Stick circuit requires a 5 volt power source to be operated, which is a reasonably safe voltage for a human being.

Figure 3.7 presents the internal circuit architecture of the Magic Stick with the different electronic elements used for the design.

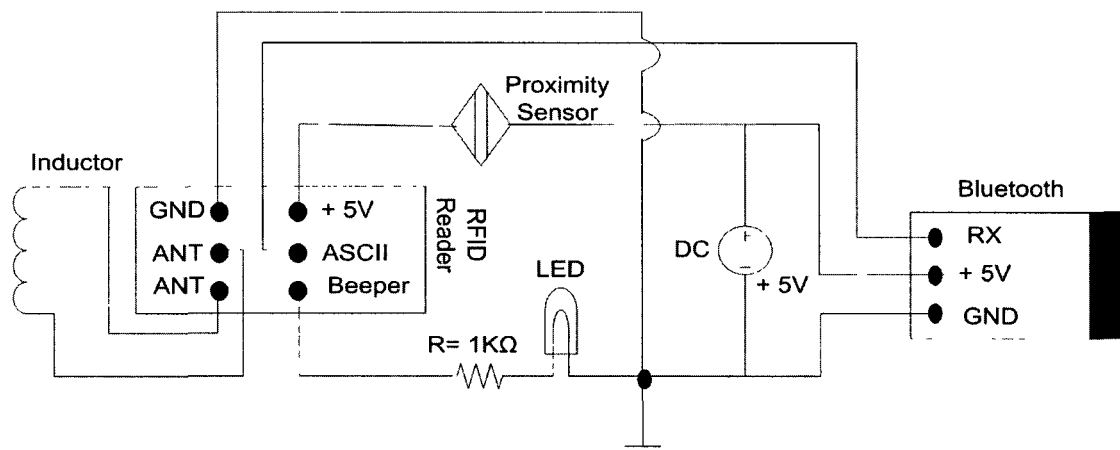


Figure 3.7 The Magic Stick Circuit Architecture [34]

Challenges of Interaction with the Magic Stick

The RFID reader can read from a maximum distance of 12 cm only. Children are not always expected to tap on the object within an allowable reading distance from the

place where the RFID tag has been placed. To overcome as much as possible this inconvenience, we maximized the RFID reading distance by building an inductor with a determined number of turns that serves as antenna for the RFID reader. The number of inductance turns L was determined using Grover's formula [40] which is given by:

$$L = \frac{\mu}{2\pi} \cdot l \cdot \ln \frac{8 \cdot A}{l \cdot w} \quad (3.1)$$

With μ (air permeability) = $4 \cdot \pi \cdot 10^{-7}$

A = Area of the loop in square meters.

l = Perimeter of the loop in meters.

w = width of the copper trace in meters.

After adding the antenna, the reading has improved to a maximum distance of 18 cm, which is considered a good reading distance for a considerably small RFID reader of low frequency (125 KHz).

Interaction with the Magic Stick

In order to clarify how interaction is done with the Magic Stick, we present the flow diagram in figure 3.8. Using the Magic Stick, the system accepts a 12 bit identification code from the RFID tag being tapped. Once a user taps on the object being tagged, the RFID reader detects that code and sends it wirelessly to the system's software running on the computer. The system provides the user with a graphical user interface where he can customize his own mapping for each tag ID. For instance, the guardian might place a RFID tag inside a bear doll and map the tag's ID with the word "Bear". In this fashion, for an object identification scenario for example, the system won't limit the number of objects that could be tagged. In other word, each time the parents or teachers think that children have already learned a certain number of entities or objects, they could simply change the mapping and place the tags inside other entities.

When the mapping becomes known, the system provides appropriate output according to the different gaming/interaction scenarios we developed. The flow of the remaining steps is self explanatory.

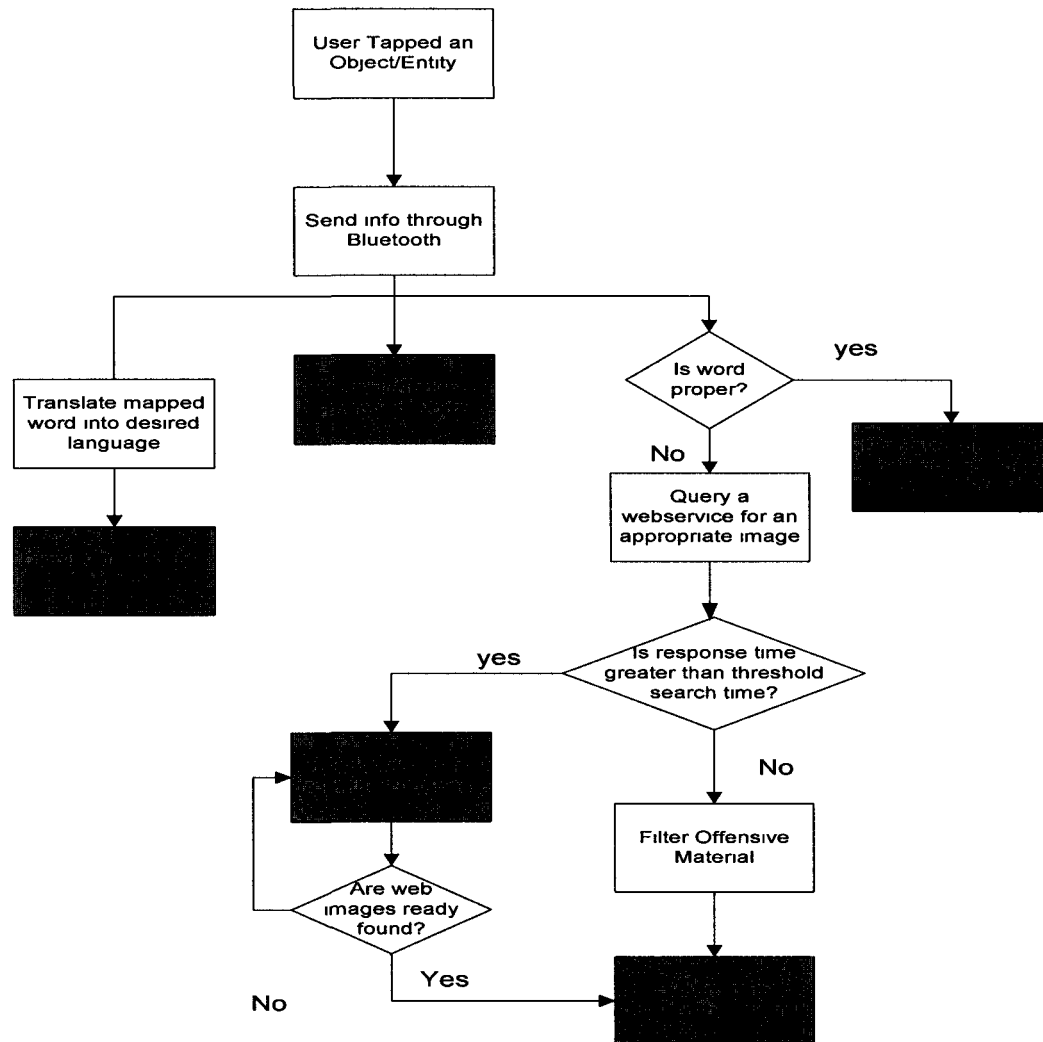


Figure 3.8 The Magic Stick interaction flowchart

3.6 Multimodal Output Selection

We have allowed the possibility of incorporating different types of media to generate appropriate output. According to the guardian's choice the different media could be displayed either all together or individually. The idea behind allowing the guardian to select his desired media output is to give him more control over the types of media that he/she thinks it is most appropriate for the children depending on their context and the scenario of the game or even depending on what he/she would like children to be exposed to.

For instance, if a mother is telling her young child a story in his bed, she can use the Magic Stick to provide audio/visual output corresponding to the objects/entities of the story. However, it would not be possible to show visual output on a screen while the child is lying in his bed unless there is a TV attached to the ceiling. Hence, the mother might only select the audio output and take advantage from either the storytelling capability of the system (stories she has already saved in the system) or the name uttering capability of the system every time she taps on a picture in the story book. Another scenario might occur in a big classroom where only one small computer screen is available. The teacher might only select the audio type of output for an object identification game since it will not be possible for all children to look at that small screen from a far distance.

3.7 System's Scenarios

3.7.1 Book Game

The first way that very young children use computers is on their mother's lap [43]. The Book game adopts this concept and allows very young children to get the benefits of edutainment while on their parents' laps. The main idea of the game is that whenever

the name of a character or object of a story book is being uttered or tapped by the Magic Stick, an appropriate media concerning that character is displayed.

Components: For Magic Stick technique, a RFID tagged story book for the young children that contains different elements and characters is provided to the guardian. For the voice-based technique, a microphone for speech input is required for all games.

Game Scenarios: In order to start the game, the guardian must first choose which type of interaction interface he/she would like to use. Considering the Magic Stick technique, a mother, while telling a story about the elements in the book, tap on a picture and get in response different pictures of the character/object, a voice uttering the name of that character/object, and a text that shows the spelling of the name. For example, she might tell the child “the bear was sleeping” and at the same time taps on the bear’s picture. The system in this case will show different images of the bear and utter the word “bear” loudly. This way the child will be able to consolidate better the name and shape of a bear and might even be able to visualize the bear when she read bear at a later time.

On the other hand, using the voice-based technique, the guardian or child will be encouraged to spell these names instead, and get the same response as in the magic stick case. For instance, if the child spells the word “Bear” as “Bee”, the system will identify that the word is “Bear” through the mother’s mappings in the system, and get in response a voice saying “Bear” with the rest of the media. In this case, spelling the name in the correct manner will help the child improve his speaking skills and pronunciation. Figure 3.9 presents a picture the GUI of the Book game and the output display when a bear picture was tapped.

Game Options: The system is also associated with a video option. Young children’s literacy skills benefit from having storybooks read repeatedly to them by a computer [44]. With such option, guardians can save short story clips that they would like their

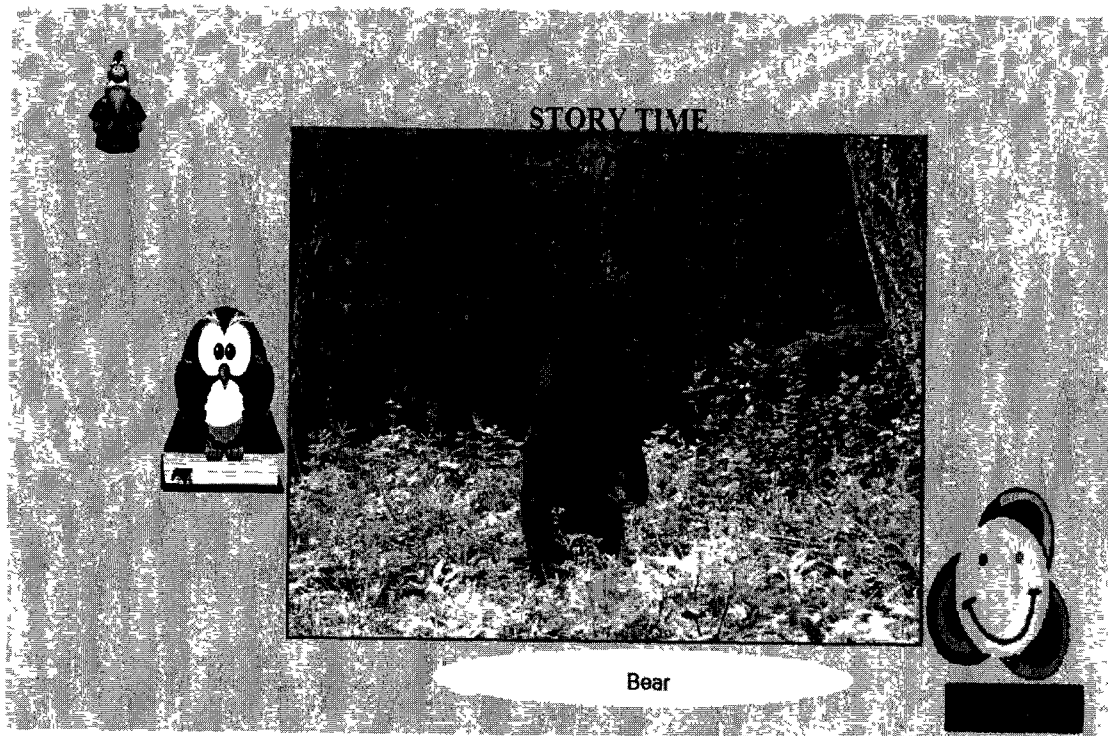


Figure 3.9 The GUI of the Book game

children to listen to. Then they can map each clip to an appropriate name of a character/object. As a result, the children will be able to hear stories and see videos every time they or their guardians utter a name or tap on a picture of the story book and for as many times as desired.

3.7.2 Learning Alphabet Game

As the name of the game might reveal, the Learning Alphabet game aims to teach children who are at early learning stages the reading and writing of the alphabets. Children can write any word and get in response pictures, voice, and text translated into a custom language regarding the word they have written.

Components Associated: For Magic Stick interaction, we have tagged the English alphabet letters using 26 RFID tags as shown in figure 3.10. We also used three other RDIF tags for performing extra operations as per the following:

- **“Bye bye”** tag that allows the user to exist the game.
- **“Show Image/Clear Screen”** tag that allows the user to request images about the word he writes or if tapped twice to clear the screen from existing images to play again.
- **“Clear”** tag that allows the user to clear a letter in case of a writing error (similar to a backspace in a regular keyboard)

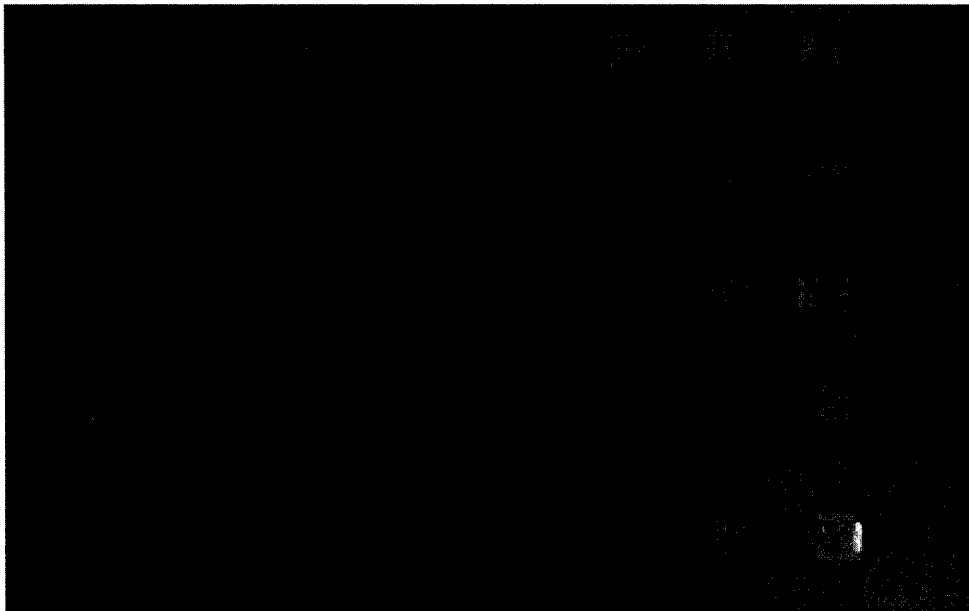


Figure 3.10 The RFID tagged keyboard

Game Scenarios: Considering the Magic Stick, One or two children can either write alone or collaborate in writing a desired word. Assuming a classroom setting, the teacher might ask two children to write the word “Apple”. Each child at a time can write one letter by tapping over the desired tagged letter. For instance, child 1 taps the

letter “A”, and then his classmate child 2 taps the letter “P”, and then again child 1 taps the letter “P” and so on until the word has been fully written. Or if desired, one child might actually write “App” and then his classmate writes the rest of the word. Each time a letter is tapped, children will see the picture of the letter being tapped, and their avatars will spell the letter for them. Eventually, when a child taps on the Show Image/Clear Screen tag, a set of pictures regarding that word will be displayed, an avatar of the child who tapped the Show image button will utter the word, and a French translated text of that word will appear in the black board. Figure 3.11 shows the GUI of the Learning Alphabet game with the different media associated.

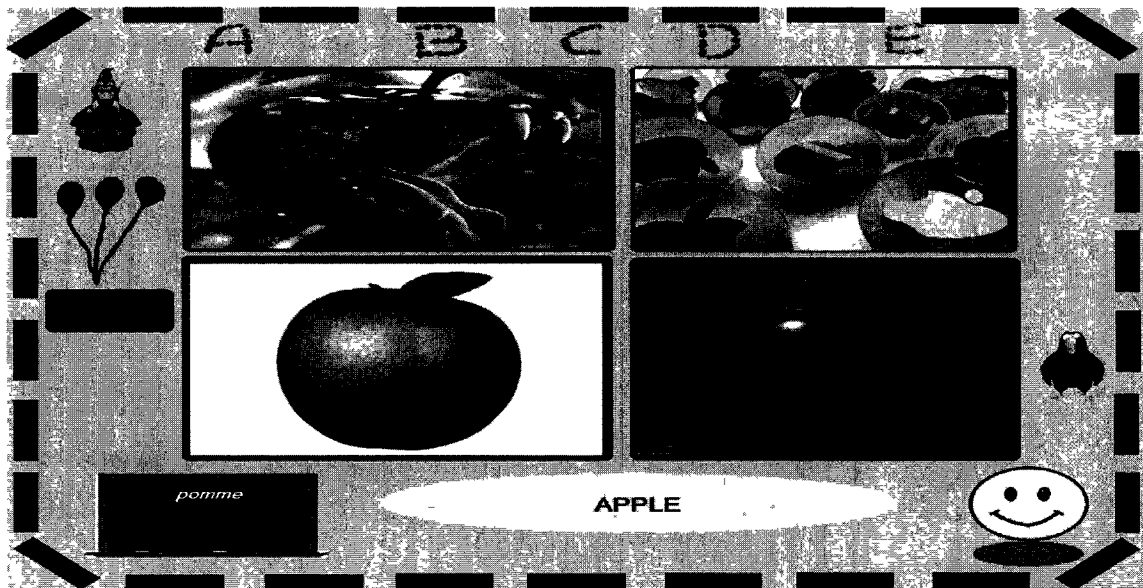


Figure 3.11 The GUI of the Alphabet Learning Game

In the case of a voice-based technique the system will accept the voice input of the tags and follow the same procedures. For instance, child 1 will utter the letter “A” while child 2 utters the letter “P” and so on. In case of an error, the child has to say clear for instance to clear a letter. When done writing, a child has to utter the word “Show Images” in order to request the pictures from the web.

3.7.3 Object Identification Game

The Object Identification game aims at familiarizing children with new objects/entities that exist in their environment. It aims also at enhancing their knowledge about entities and objects they know by showing them different colors and shapes of those entities and shapes.

Components Associated: Some RFID tagged dolls of animals, birds and insects.

Game Scenarios: Considering the Magic Stick as an interaction interface, the teacher might ask one or two children to look for four different types of birds among the dolls in the classroom. The children can either alone or in collaboration in a group of 2 solve the quiz. For instance, child 1 might identify 3 types of birds while child 1 identifies the fourth one. Similarly to the previous games, each time an object/entity is tapped by a child, some pictures will be shown while the child's avatar utters its name. A French translation is provided at the end in the black board. Figure 3.12 shows the GUI of the Object Identification game with the different types of media associated.

Now considering the voice technique as an interaction interface, children will have the opportunity to learn about new colors or shapes of objects they already know. Since children are expected to utter themselves the name of the entities, we expect them to have a small knowledge about them, and therefore we could widen their knowledge by showing them their various shapes or colors. For instance, a child might utter the word apple, and get in response a set of pictures showing different apples in different colors.

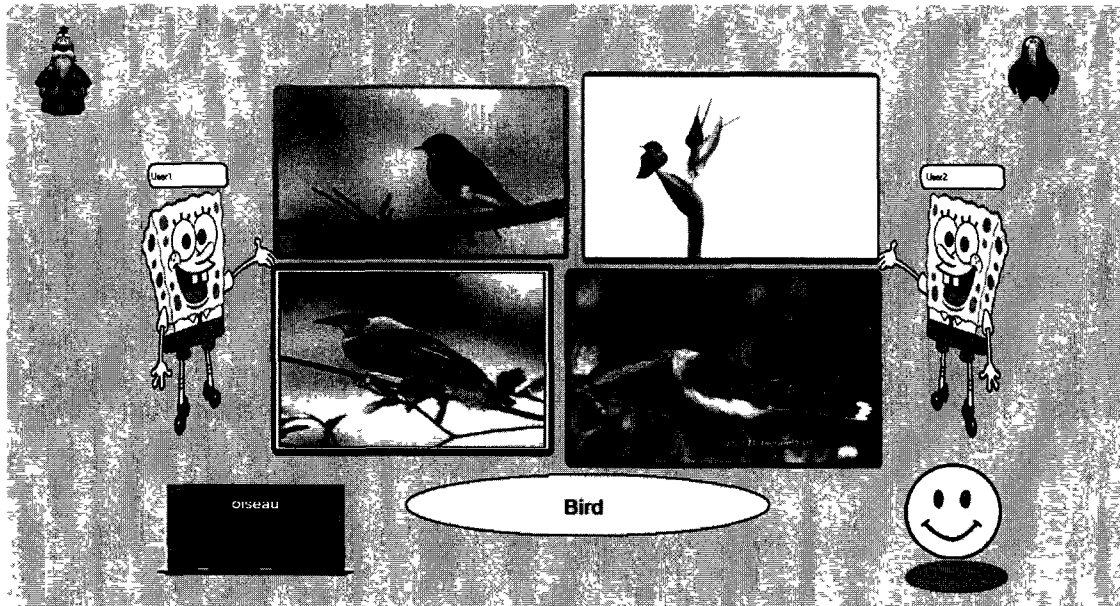


Figure 3.12 The GUI of the Object Identification Game

3.8 Summary

In this Chapter, we have set the requirements of our proposed system. The overall system architecture was discussed with the different services and modules incorporated. In addition, we explained how we met our requirements and gave an overview about the three different games associated with the system that satisfy those requirements.

Chapter 4

Implementation

This chapter gives details of the component and interaction diagram of the developed system prototype (Section 4.1.), provides description of the devices and APIs used (Section 4.2) and shows the graphical user interface of the prototype (Section 4.3).

4.1 Component and Interaction Diagram

This section describes the structural and behavioral designs of the Magic System. For instance, Figure 4.1 presents the software component diagram that demonstrates the components associated with the architecture discussed in Section 3.3 of Chapter 3.

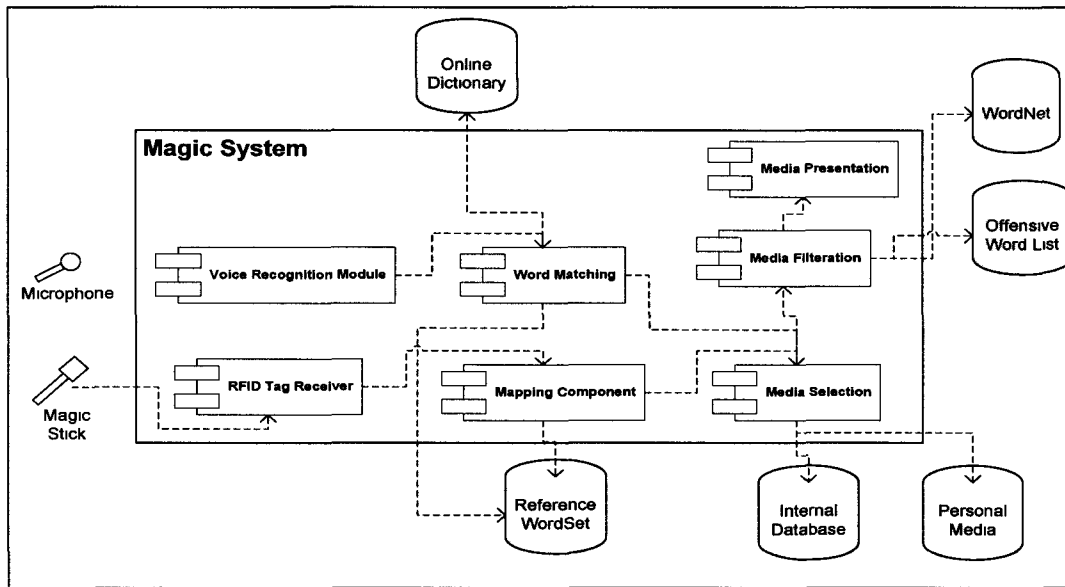


Figure 4.1 The component diagram of the Magic System

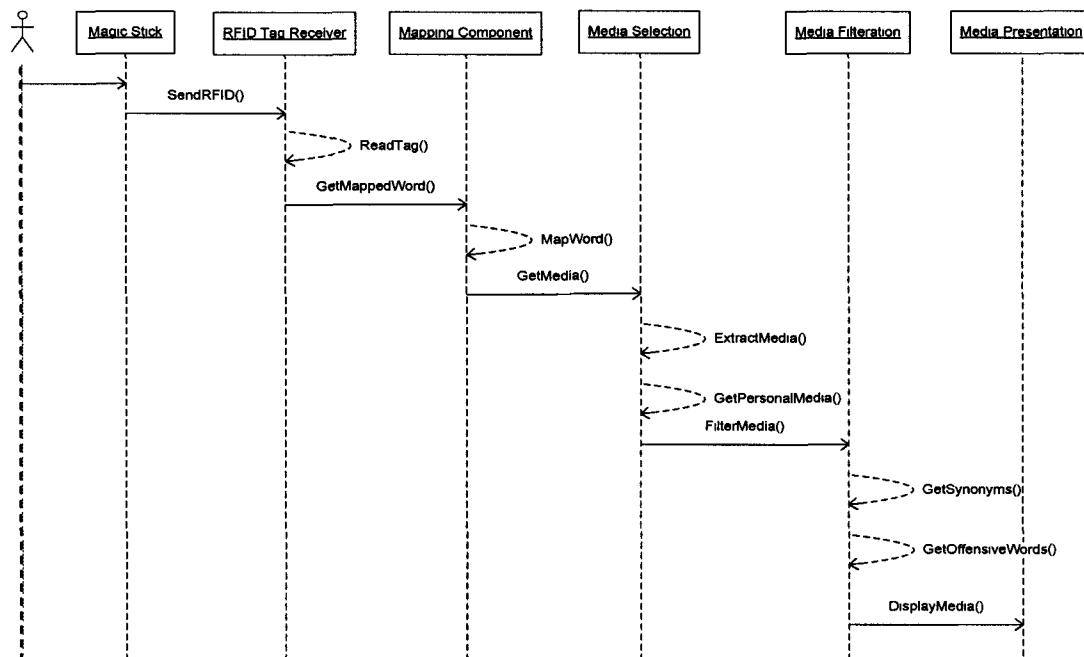


Figure 4.2 The interaction diagram of the system with the Magic Stick technique

Interaction diagram with the Magic Stick: The interaction diagram of the Magic Stick is presented in Figure 4.2. The user triggers the system by simply tapping over a RFID tag. Once a tap occurs, the RFID tag ID is sent to the RFID tag receiver, which is in this case a dongle attached to the computer that reads that ID and passes it to the Mapping Component. Then at the Mapping component stage, the mapping of the tag ID with the custom word or term is done and a request is sent to the Media Selection module to extract an image either from a personal repository that was defined by the guardian or from a webservice. After this stage, the Media filtration will filter the extracted images by examining the keywords associated with each of them. This will be done through checking if any keyword matches an offensive word and by calculating the hyponym distances as was explained in Section 3.3 of Chapter 3. Finally, after the filtering is done, the images are displayed by the system.

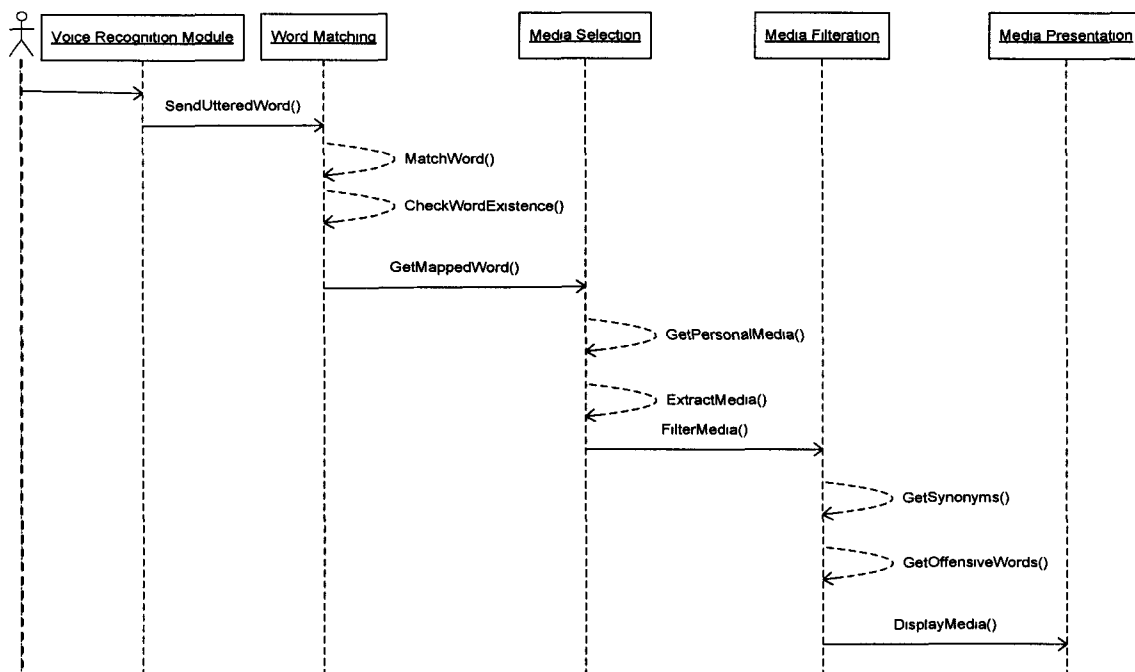


Figure 4.3 The interaction diagram of the system with the voice-based technique

Interaction diagram with the voice-based technique: The interaction diagram of the voice-based technique is presented in Figure 4.3. The user triggers the system by uttering a word that will be detected by the Voice Recognition Module. After the word detection occurs, the detected word is sent to the Word Matching module that will find an appropriate match. The rest of the sequence is similar to the steps discussed previously.

4.2 Devices and APIs

In this section, we give the specific of the different devices and the different APIs we used to develop the whole system, as below:

Magic Stick Elements: The actual design of the Magic Stick is presented in Figure 4.4. It had been implemented using a Core-12 RFID short range reader module that reads with a 125 KHz frequency. The RFID module was attached to a BlueSmirf Gold Bluetooth chip that supports a baud rate of 9600-115200 bps. For tagging objects, we have used a 30 mm Disc tags- Unique/EM4102. An indicator LED that indicates ID tag detection was attached to the RFID reader. The stick was empowered through a 9 V battery connected to a voltage regulator that supplies + 5 V to the whole circuit. A basic proximity sensor attached to a transistor played the role of an automatic switch. A manual switch was also added in order to allow for a complete shutdown of the Magic Stick. The Magic Stick components were actually mounted on a light weight piece of wood. Figure 4.5 shows the component diagram of the Magic Stick.

Voice Input Device: We have used Labtec microphone for input voice detection. The identification of user speech was done through the Microsoft Speech Engine service.

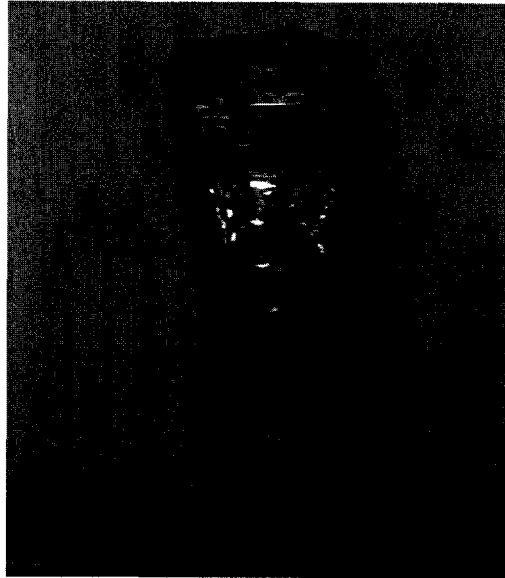


Figure 4.4 The Magic Stick

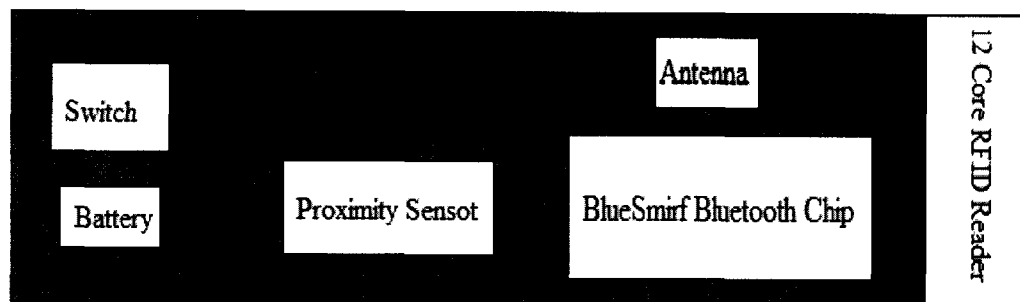


Figure 4.5 The Component diagram of the Magic Stick

Software Module and APIs: The proposed system has been implemented using Visual Studio 2005 (C# language) on a Windows XP platform. A multi-threading approach had been adopted during the implementation, which was needed to execute some of the system's functionalities over multiple threads. For example, the voice recognition service was executed on one thread while the image search service was executed on

another thread. Since the fetching process of the images might take some time, sending multiple requests to the search module while the fetching process is not complete might impose an overload on the module. For this reason, the voice recognition thread was halted once a word has been detected and the request was sent to the search service to reduce the overload on the search module. However, the voice recognition thread is recreated as soon as the extracted image is displayed on the screen to make the system ready to take further inputs. The system software used different APIs for the different functionalities that the system provides which can be states as follows:

- **Google Translate API:** The Google Translate API [47] allows translation from any language into any other language that the API supports. In order to use it, we have to first specify two languages. The system's default language (e.g English) and the system's output translated language (e.g French). Then we simply provide a string (word) to the Google translate API, and get in response another string (translated word). Figure 4.6 shows sample syntax that demonstrates how this could be achieved: We have associated translation from English to French in our system; however, the guardian or the user can pick the languages of choice through the system's GUI.

```
String WordToBeTranslated, TranslatedWord;  
GoogleTranslatorAPI.GoogleTransAPI GTAPI = new  
GoogleTransAPI(GoogleTransAPI.GetEnglishString(),  
GoogleTransAPI.GetFrenchString());  
TranslatedWord=GTAPI.TranslateText(WordToBeTranslated);
```

Figure 4.6 A Google Translate API send request/receive response sample syntax

- **Microsoft Agent API:** The Microsoft Agent API [48] is a technology that employs a text-to-speech engine and some animated characters that spell the text to the users. We first choose the names of the animated characters to be used.

Then we provide a string word or text to the Agent API and get in response a speech or a spelling of that string. The movements of the agents can be controlled by changing the parameters of their X-Y coordinates. We have employed two agents: Merlin the magician and Peedy the bird. Figure 4.7 demonstrates how the Microsoft Agent API is used.

```
String WordToSpeak;  
axAgent1.Characters.Load("Merlin", "merlin.acs") //Load Agent  
speaker = axAgent1.Characters["merlin"];  
private AgentObjects.IAgentCtlCharacter speaker;  
this.speaker.Show(0); // Start the animated agent from origin  
this.speaker.MoveTo(100, 100, 1); //move the agent within these coordinates  
this.speaker.Speak(WordToSpeak, "");
```

Figure 4.7 A Microsoft Agent API send request/receive response sample syntax

- **Web Images API:** We have created a Meta search API for extracting images from Yahoo and Flickr services using the respective search facilities. First, we specify the maximum number of images to be returned by the search service for a particular request. Then we provide a word as input to the search API and get a set of images as response. For the same request made at different times we obtain different images from this service, which provides a dynamic nature of output. This is due to the fact that the image base of these two providers is constantly being updated with the increase of web resources and/or contribution from general people.
- **WordNet API:** This is a lexical database of English language [42], which has been used to realize the functionalities of some components. For example, it is used in our system to help the voice recognition service to identify the synonyms of the words that children utter. Returning some set of synonyms by

```

/* WordRequested is a String type variable that contains the user's requested word,
accessible using the getUserDefinedWord() method.. We generate all the senses
(synonyms) of this word and store them in the ArrayList Variable MySense for
further processing*/
String WordRequested = getUserDefinedWord()
Index idx
MySense = new ArrayList()
/* Obtain the reference of the wordRequested in the Dictionary and create the
similarity list of words. POS refers to the Parts of Speech of WordRequested */
Indexes ixs = new Indexes(WordRequested, pos)
/*Loop until there are no similar words to be processed for WordRequested*/
While ((idx = ixs.next()) != null)
Begin
    for (int sens = 0; sens < idx.off.Length; sens++)
        Begin
            /* For each senses, we load the word and add it to the MySense
List*/

            SynSet cursyn = new SynSet(idx, sens, this)
            wordsFrom(cursyn)
            MySense.Add(cursyn)

        End /*For loop*/

```

Figure 4.8 A WordNet API algorithm for requesting words synonyms

the WordNet API means that a word uttered has a meaning in the English language and could be sent to the image search engine. Also, it is used to help the filtering service identify whether an image was offensive or not through providing the synonyms of the image keywords that will be used to calculate the hyponym distance among the image keywords and those synonyms. These two services were explained in details in Section 3.3 of Chapter 3. In order to use the WordNet service, we provide the API with the word (string) we want to find its synonym and get in response an array of strings where

each string is a synonym of that word. Figure 4.8 shows a WordNet API algorithm that clarifies how to send a request and get its corresponding response.

4.2 Graphical User Interface

An important aspect in the design of the system is the GUI that should be easy to use. We have designed 3 different interfaces for three different games and one main interface where the settings of those games can be customized. It is worth mentioning that if the guardian wants the children to see the visual output, only then the UI will be needed. Otherwise, the system can be configured to provide audio output and in that case the use of UI will be ignored. As can be seen in figure 4.9, the interfaces of the games 1, 2, and 3 don't look like the traditional windows form (as interface 1). We have used the C# Custom Shape Form Region [50] that allows us to draw irregular form shapes by changing a single Form property called Region. Avoiding traditional windows forms will give children the feeling that they are not playing on a regular computer and enhance their interactions with the system because of the various drawings and colors that are incorporated.

Through the main user interface (number 4), the guardian (or children with some computer skills) can change the settings of each of the games. For instance, he/she can select the type of media to be associated with each game such as video, audio, images ...etc. Moreover, he/she can select the input interaction interface (Magic Stick, Speech) and modify the related Word and ID Mapping information related to each technique. Figure 4.10 presents a screenshot of the ID Mapping interface where the guardian can associate with each RFID tag an appropriate word map, object/entity type, description and an image location. The number of mappings is not limited and more columns or rows could be added or removed depending on the scenario of the game, or the scenario that the guardian himself may like to follow.

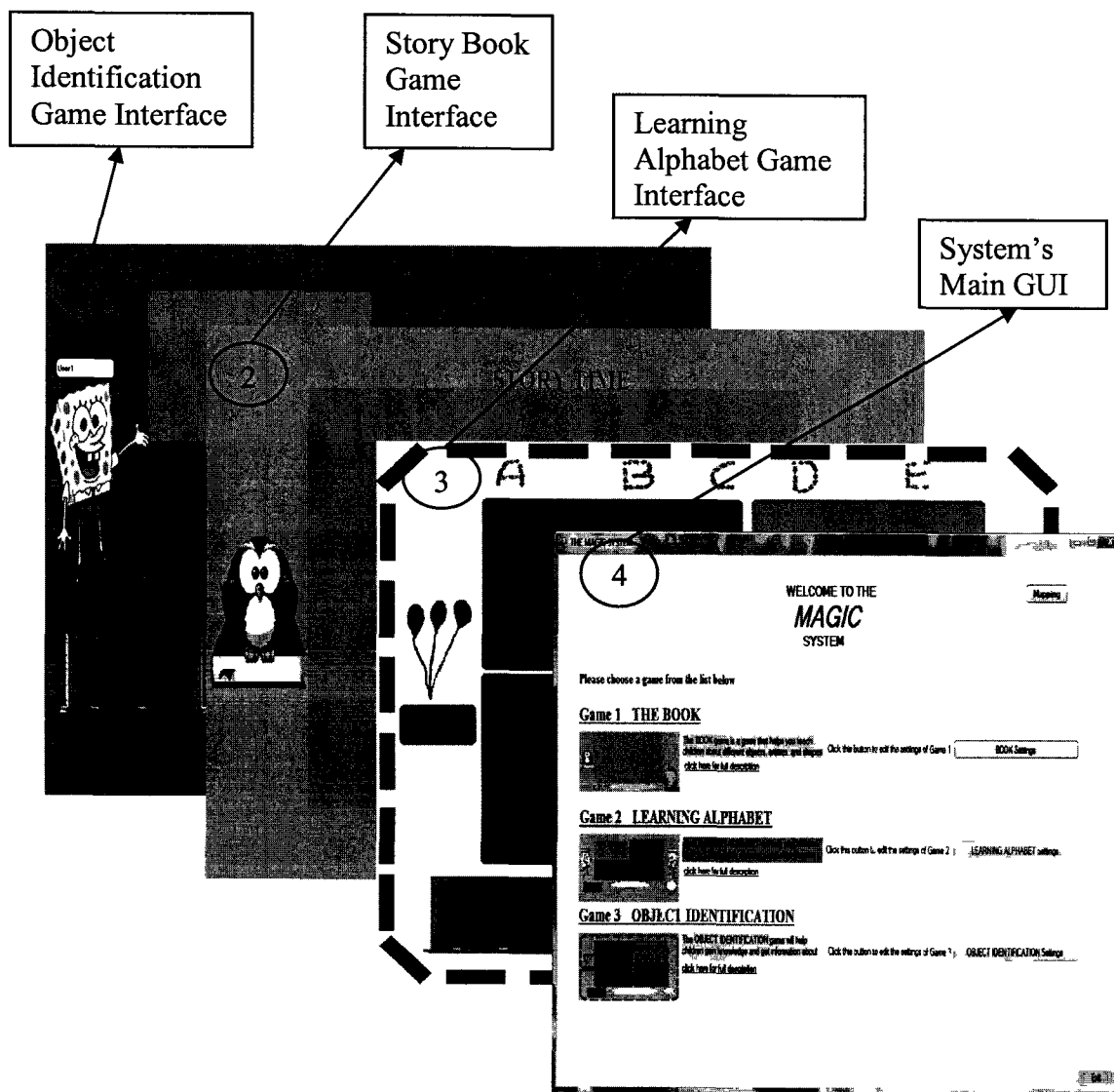


Figure 4.9 The GUIs associated with the system

Tag Mapping

Add/Update Mapping Information

Exit

TagID	Word Map	Object/Entity Type	Description	ImageLocation
01023875551B	Z	animal	zebra	C:\pictures\DSC001
010238AA0392	Y	Fruit	Yellow	C:\pictures\DSC002
01068CAEA98C	T	Insect	Worm	C:\pictures\DSC034
0104FF15AB44	Show Image/Cle	Sign	Stop	C:\pictures\DSC005
01068CB3457D	S	Insect	Spider	C:\pictures\DSC007
0103C7F886BB	K	Fruit	Plum	C:\pictures\DSC008
01057027EFBC	P	Fruit	Pear	C:\pictures\DSC023
01068CBD7E48	O	Fruit	Olives	C:\pictures\DSC67
01068C9EA9BC	Q	Insect	Mosquito	C:\pictures\DSC28
0104FEF080BB	M	Animal	Monkey	C:\pictures\DSC009
010238A6A835	J	Fruit	melon	C:\pictures\DSC475
01068CACBF98	L	Animal	Lion	C:\pictures\DSC357
01068CACE0C7	G	Fruit	Grapes	C:\pictures\DSC57
01068C9E485D	F	Animal	Frog	C:\pictures\DSC101
0104449029F8	R	Insect	Fly	C:\pictures\DSC341

Add More Columns Save Mapping

Figure 4.10 The GUI of the ID Mapping table

Chapter 5

Evaluation

In this chapter, we present the evaluation of the proposed edutainment system. In particular, we evaluate the system based on the three gaming scenarios as presented in Section 3.7. The evaluation was performed by considering children of different age group to interact with these games. Accordingly, in section 5.1 we show the evaluation of the Book game scenario among preschool children, ages 2-4, using both the voice-based and the Magic Stick interaction techniques. In section 5.2, we test our Alphabet Learning and Object Identification games among kindergarteners using the Magic Stick interaction technique.

5.1 The Book Game Evaluation among Preschoolers

We have tested the Book game among preschool children ages 2-4 years old using both the voice-based and Magic Stick interaction techniques. We explain these in the following:

5.1.1 Voice-based Interaction technique evaluation

We start our evaluation for the Book scenario by allowing children to interact with the system through the voice-based interaction technique.

- **Experimental setup**

To evaluate the voice-based technique among children, we have visited 5 families at their homes to test with their children. We have a laptop for running the system's software and display the appropriate media output. We have connected also a powerful Labtec microphone to the computer for speech detection and we placed it on a table facing the screen.

- **Experimental procedure**

We conducted our qualitative evaluations with 5 children, 2 girls and 3 boys, ages 2-4. We have chosen this age range since normally children of 2 years and older can mumble words of things they see. Children were sitting on their mothers' laps, and the experiments were conducted in 5 different days within a 2 week period. Each experiment session has taken approximately 30 minutes and the sessions were videotaped for further analysis.

Before the day of the first experiment, we asked children's mothers to provide us with a set of words that their children could utter the most. The reason behind that was to create a list of common terms that all children could utter so that we could analyze our results based on those terms only. In addition, we wanted to know previously these terms so that we could prepare the appropriate story book that contains pictures related to those terms. We were able to choose 7 common terms that all of the children could utter. We then defined those terms in the first column of the mapping table through the

GUI that contains information regarding the words children can say, and defined if any word was proper or not. The way of defining the mapping table was explained in Section 3.3 of Chapter 3.

On the day of the experiment, we have explained for each mother how the system works and demonstrated how to use the GUI that allows them to define their children's style of utterance. In addition, we have asked each mother to train her child before the experiment on spelling the 7 terms while showing him/her the pictures in the storybook by allotting a 10 minute period for each mother before the beginning of each experiment.

- **Test Cases**

We have tested the voice-based interaction by playing the Book game. The scenario required the mother to show her child the book's pictures of the objects or entities and encourage her child to spell the names of these objects and entities. The objective of testing with the Book game was to test the system's ability of identifying children's words and displaying appropriate images.

- **Results**

To test the ability of the system of correctly identifying the words uttered by the children and display appropriate media we have asked each mother to encourage her child to repeat the 7 words of the list for 3 times. The word detection of the young children was a very challenging problem because of their pronunciations which might be different each time they spell and their mood which might not be always stable. For 2 of the children, repeating the words was not a big problem; however, for the rest of the children it was hard to force them to say all the terms for many times. For instance, some of the children were only repeating 3 or 4 words again and again and refusing to spell the rest. Others were saying mumbling different words that don't exist in our list.

Since we wanted each of the 5 children to utter the 7 words of the list for three times, the total number of words that we expected them to utter was 105. However, some children didn't spell all of the 7 words from the list, so the total number of words came down to 75. The number 75 is actually the total number of words that the children have spelled from the list. We didn't take into consideration the other terms they said (words out of the list) since we wanted to test the detection capabilities of words already defined in the system. Each child has spelled an average of 5 words for 3 times. We have examined the ability of the system to identify children's utterance by considering the number of times a child has to utter a word before it gets detected by the system. In other words, we measured the number of words that the system has successfully detected from the first attempt a child has said a word, the second, and the third one. Table 5.1 presents the number of successful word detected for each of the first three attempts.

Attempt number	Total number of words detected	Total number of words uttered (out of 75)
First Attempt	18	18
Second Attempt	6	12
Third Attempt	3	6

Table 5.1 Results obtained from children's utterance of 75 words

The results obtained in table 3.1 can be explained as follows:

The **total number of words detected** presents the number of successful words that the system detected for each of the three attempts. The system was able to detect 18 words from the first attempt, 12 words from the second, and 3 words from the third.

The **total number of spellings** presents the equivalent number of words that the children have uttered for each of the successful attempts. For instance, if a word has been successfully detected from the second trial, this would mean that it required the child to spell it 2 times. Hence, 6 words detected from the second attempt is equivalent to 12 words spelled out of the 75 words.

- **Discussion of the results**

The results in table 5.1 shows that only 36 words (18 +12+6) out of the 75 word have been successfully detected while the remaining 22 were undetected by the system. The chart of figure 5.1 presents a graphical sketch of the percentage of word identification success among the 75 words that have been uttered by the 5 children. As can be concluded from the chart, almost 24% of the words were correctly identified from the first time (attempt) children have spelled a word, while 16% from the second time, and 8% from the last. This would lead to a 48% detection error which is not an acceptable error rate. Upon these results, we have used the alternative interaction approach (the Magic Stick) to eliminate the drawbacks of the voice-based technique.

Problems Encountered: The speech engine deactivates every time a word is detected and then reactivate when an output is displayed as was explained in Section 3.5 of Chapter 3. Since mothers will be speaking with their children, the system will always detect first the words of the mothers and might not be always ready to detect what the children are saying. For this reason, we have used a switch for turning manually on and off the speaker during the experiment. In this manner, the mother can switch off the microphone every time she is talking, and then switch it on when she is done so that only the words her child might say are detected.

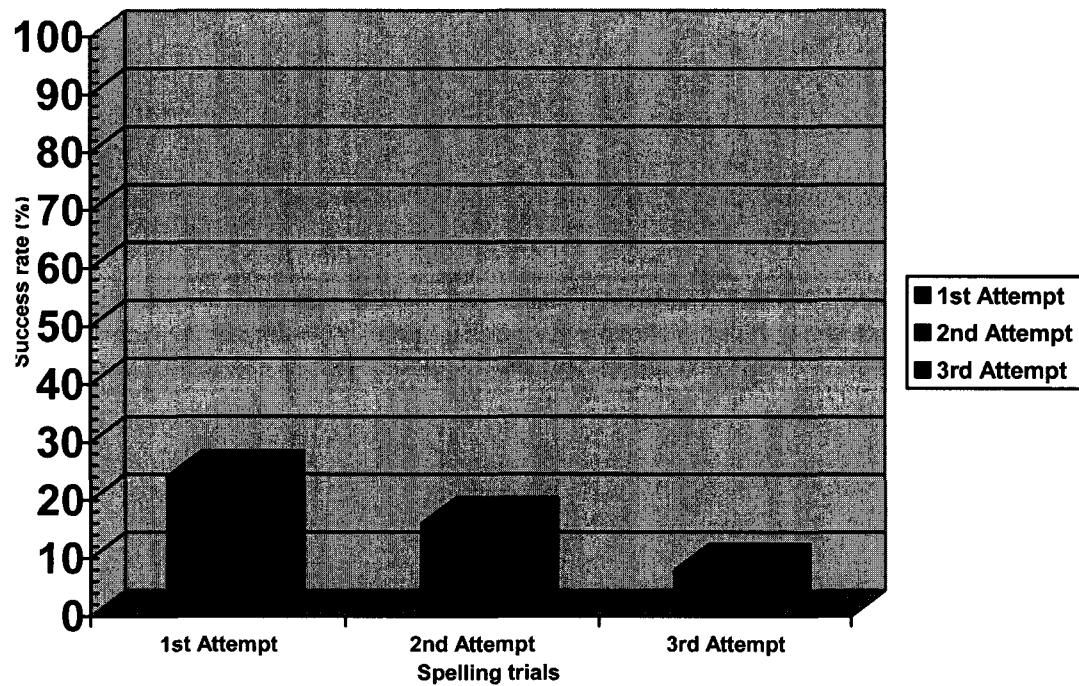


Figure 5.1 The success rate of each of the three trials

5.1.2 Magic Stick Interaction technique evaluation among very young children

After the inconveniences children have encountered with the voice-based technique, we have evaluated our system with the Magic Stick technique.

- **Procedure**

In order to test the Magic Stick with the very young children, we have used a story book that contains many pictures of objects and entities. We used a thick page story book (pages were made of hard carton) so that tagging the pictures would be possible. We attached for each picture in the book a RFID tag and modified the ID mapping tables correspondingly so that each picture has an appropriate name or description related to it. A Bluetooth dongle was attached to a desktop computer and the results were show on a 19 inch screen.

We conducted the experiment with the same 5 children and mothers who performed the previous experiment. The experiments were performed at the subjects' places in 5 different sessions over a week. Each session was videotaped for further analysis.

- **Test Cases**

We provided each mother with the story book and asked her to allow her child to tap on the pictures of the book. Children sat on their mothers' laps and they were both facing the screen of the computer. The objectives of the experiment were the following:

- a) Observe children's interaction with the system when using the Magic Stick.
- b) Test the Magic Stick's ID detection's efficiency.
- c) Test the system's image display efficiency by measuring the time it requires to retrieve the appropriate pictures from the web.

- **Results**

We have tested the system using the Magic Stick based by examining the following:

Children's Interaction: Children grasped the idea of the Magic Stick very quickly and started tapping over and over again on the different pictures of the book. The spellings of the animated agent of the names have urged the children to repeat what he said. For instance, when Elena heard the agent saying the word "Apple", she has immediately repeated the word and tapped again on the apple's picture. Children seemed to be excited and happy as most of the mothers have told us at the end of the sessions.

Tag Detection: The Magic Stick has detected successfully the identifications of the RFID tags that were attached to the pictures. This was obvious through the output that

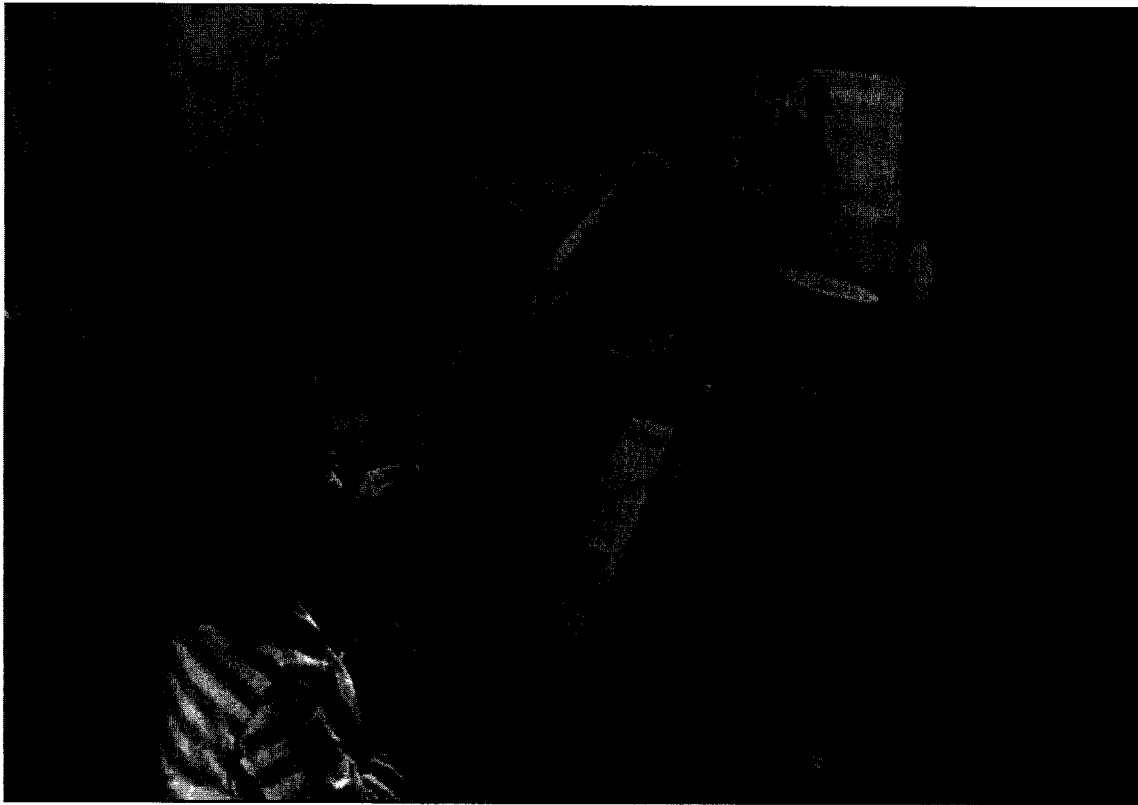


Figure 5.2 Mary tapping with the Magic Stick on a duck's picture

was displayed almost every time a child had tapped over a picture. Figure 5.2 shows the displayed output when Mary has tapped over a picture of a duck. In order to have a better estimation of the efficiency of the Magic Stick's ID detection, we have measured the number of times the Magic Stick has successfully identified a tag ID from the number of taps made by the children. Table 5.2 presents the results obtained at the end of the 5 sessions.

Children have actually tapped a total of 206 times on the pictures of the book. The Magic Stick has detected 138 tag IDs from the first tap, 14 times from the second, 8 times from the third, and 4 times from the fourth. None of the children has actually required more than 4 attempts (taps) in order to get an output by the system. The reason

Number of Attempts	Number of tag ID detected	Equivalent tap number (out of 206)
First Attempt	138	138
Second Attempt	14	28
Third Attempt	8	24
Fourth Attempt	4	16

Table 5.2 Results observed after the 5 sessions

why some children required more than one tap to get a response might depend on their ages. The more mature a child is, the better he is expected to tap correctly on a figure. Figure 5.4 shows a chart that indicates the rate of success of the Magic Stick in identifying the tag IDs from the first, second, third, and fourth attempt made by the children. As can be concluded from the chart, the Magic Stick has successfully detected about 67% the tag IDs from the first time children have tapped on a picture, 13.6 from the second time, 11.6 from the third time, and 7.8% from the fourth time.

It is also worth noting that we have also realized that in order for the Magic Stick to correctly detect the tag ID each time a child taps on a picture, there should be a time gap of at least 1.5 seconds between two consecutive tapping attempts. In fact, during this time gap the COM port of the computer will be engaged in receiving the serial data sent by the Magic Stick.

Images Response Time: To examine the efficiency of the system in displaying media, we measured the time required by the system to extract the appropriate images from Yahoo. As was explained in Section 3.5 of Chapter 3, our system displays a picture from the local database in case of a delay when fetching images from the web. We had first to deactivate this option so that we can measure only the time the system needs to

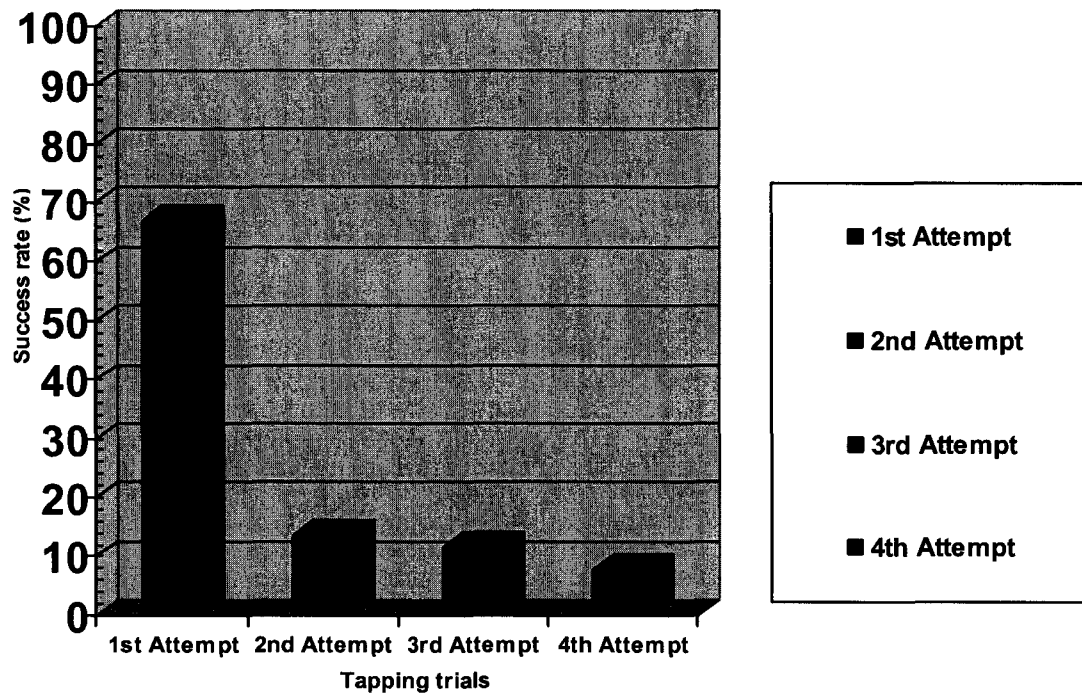


Figure 5.3 The rate of success of the 4 tapping attempts of the children

display images from the webservice. Since each child have tapped more than once on the most of the pictures of the book, we have measured the display time by finding the average time the system required to display 7 pictures related to 7 common words that were tapped during the 5 days of testing. Table 5.3 presents average the time required by the system to display an image for each of the seven common words during the 5 experimental days. This time actually includes the communication delays faced when retrieving the images. The experiments were conducted at different times during the days between 10 AM and 3:30 PM.

• Discussion of the Results

Pondering over the results, it is obvious that highest response time occurred on day 4 with Child 4 (experiment was done at 3 PM) with a value of 4.99 seconds. In addition,

Word Tapped	Child 1 Day 1 (sec.)	Child 2 Day2 (sec.)	Child 3 Day 3 (sec.)	Child 4 Day 4 (sec.)	Child 5 Day 5 (sec)
Orange	2.58	2.54	2.72	1.31	1.05
Milk	1.81	1.92	1.51	3.10	2.01
Duck	2.11	1.99	1.01	3.05	0.99
Apple	1.57	1.61	1.33	1.89	1.03
Toothbrush	1.96	1.95	2.57	2.55	1.66
Cat	2.14	2.12	4.99	2.01	2.05
Bear	1.57	1.60	2.65	1.32	2.11

Table 5.3 Average time needed for image display of the word uttered

we realize that the image response average time for the same word over 5 days is not the same. For instance, the image response average time for the word “Cat” on day 3 was higher than the average time for the same word for the rest of the days. Since the images were extracted from Yahoo, it is not possible to verify the exact reason for the long extraction times or the differences among extracting images for the same word. However, the reason might be due to one of the following:

- **Internet connection:** The internet connection plays an important role in fetching the images. A slow connection might greatly affect the output. For instance, the connection might be slower or faster depending on the time during the day. For instance, if the system requests some images during internet peak time usage such as in the afternoons, the retrieval might be slow due to large number of people using the internet at that time. This might explain why the longest extraction value of 4.99 sec occurred on day 3.
- **Filtering process:** Since Yahoo updates its databases on a regular basis, the filtering process might differ from a particular time to another. For instance, the system might not need to filter the images fetched for an orange on day 1

since none of the images tagged keywords were of an offensive type. However, on day 2, and since yahoo might have added new images of an orange with offensive tagged keywords, the system will take more time to filter the images and pick the most appropriate one (s).

- **Yahoo slowdowns:** Almost all web services perform routine maintenance that might at any time cause some slowdowns when requesting one of their services.

5.2 The Alphabet Learning and Object Identification Game's Evaluation

We have evaluated the Alphabet Learning and Object Identification games by using the Magic Stick Interaction technique only because of its simple set up and usage especially in loud environments like kindergartner classrooms.

5.2.1 The Alphabet Learning Scenario

We have started our experiment among kindergarteners by first introducing the Alphabet Learning Game using the Magic Stick.

- **Experimental setup**

To evaluate the Magic Stick among children, we have set up the system in a kindergartner classroom. We have connected a projector for media display, as well as large speaker, to a Laptop where the system's software was running. We have projected the media against the wall of the classroom and placed a table in front of that wall. On the table, we placed a green carton board that has tagged English alphabet, as well as

two Magic Sticks, as shown in figure 5.4. The alphabets and words were actually printed each on one side of the board and the tags were glued on the other part of it. In addition, we asked the teacher to place some tagged animal dolls and other small objects in the classroom.



Figure 5.4 The alphabet tagged board with a pair of Magic Sticks

- **Experimental Procedure**

We conducted qualitative evaluations with 6 groups of 2 children each, ages 4-5. Each group was given 2 Magic Sticks (One Magic Stick for each child). The evaluations were conducted in a two different scheduled sessions. Each session took approximately 60

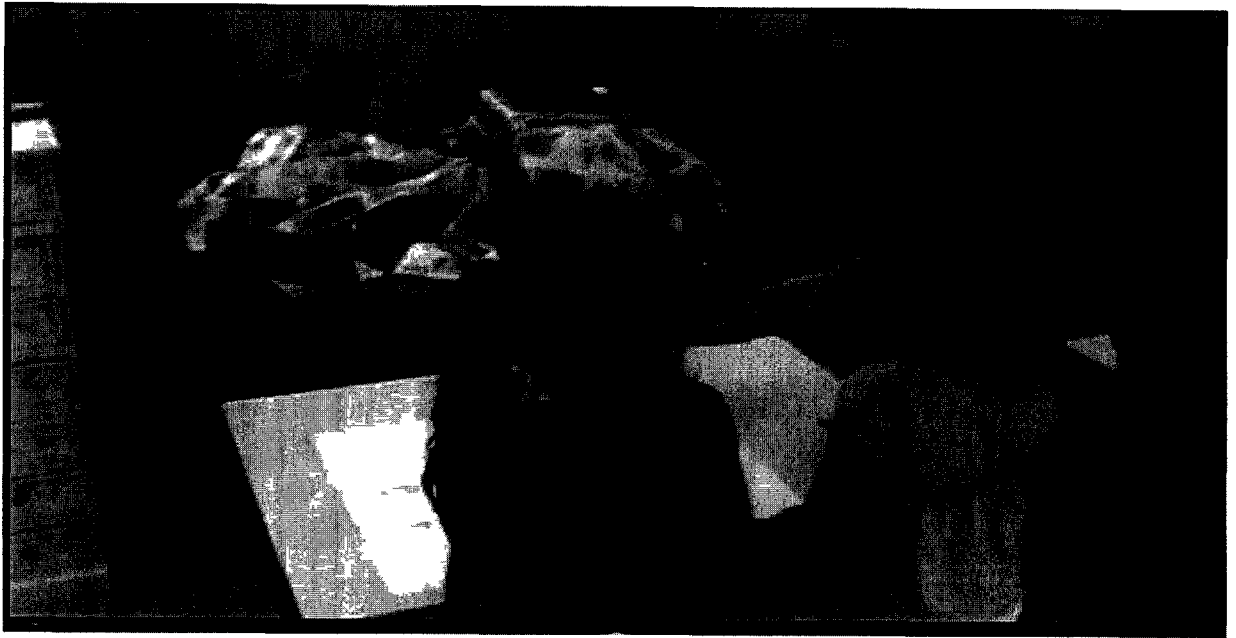


Figure 5.5 A teacher explaining the concept of the Magic Stick

minutes. Both sessions were videotaped and the interaction of the children were later observed and analyzed.

Before conducting any experiments, we first had to introduce how the Magic Stick works. We also showed the teachers how to use the Magic Stick, as well as the different functionalities of the GUI associated with the system. We stressed the importance of using the Magic Stick softly by tapping in a slow manner on the letters and on the top of the objects/entities. In addition, we explained how to modify the tag identification mapping tables.

Now for children, the teacher has presented for each group of children a small demo on how to use the Magic Stick by writing a word that returned in response 4 pictures, and a French translation of the word, as well as an audio spelling by the animated agent. Figure 5.5 shows a picture of the teacher while explaining how the Magic Stick works.

- **Test cases**

Each group of children was either asked by their teacher to write a certain word or to write a word of their choice. Each group was also asked to write a word either collaboratively (2 children writing one word together) or individually (one word for each child). The objectives of testing with the Alphabet Learning Game were the following:

- a) Observe children's interaction among each other by analyzing their discussions, writings, and queries.
- b) Check the level of simplicity of using the Magic Stick by examining the number of times children have successfully tapped on a letter, and examine the difficulties children have faced when using it.

- **Results**

We have based our result analysis based on the objectives of the Alphabet Learning game mentioned previously. Our analysis focused on the following:

Discussions: Children helped each other in solving the games through sharing ideas and discussions. For example, when writing "Apple", one child told his classmate that the word "Apple" requires 2 P letters and not only one. Another child told his classmate "can I write pig alone and then you can write yours after", and then the other child said "Ok". They also helped each other finding the alphabet letters on the board when someone had a difficulty finding a certain letter as shown in figure 5.6.

Writings: The words that children have written varied between names of cartoon characters they watch on TV, names of fruits, vegetables, animals and food in general. For instance, Ashley has written "Pizza", while Sara wrote "Cat". Figure 5.7 shows the

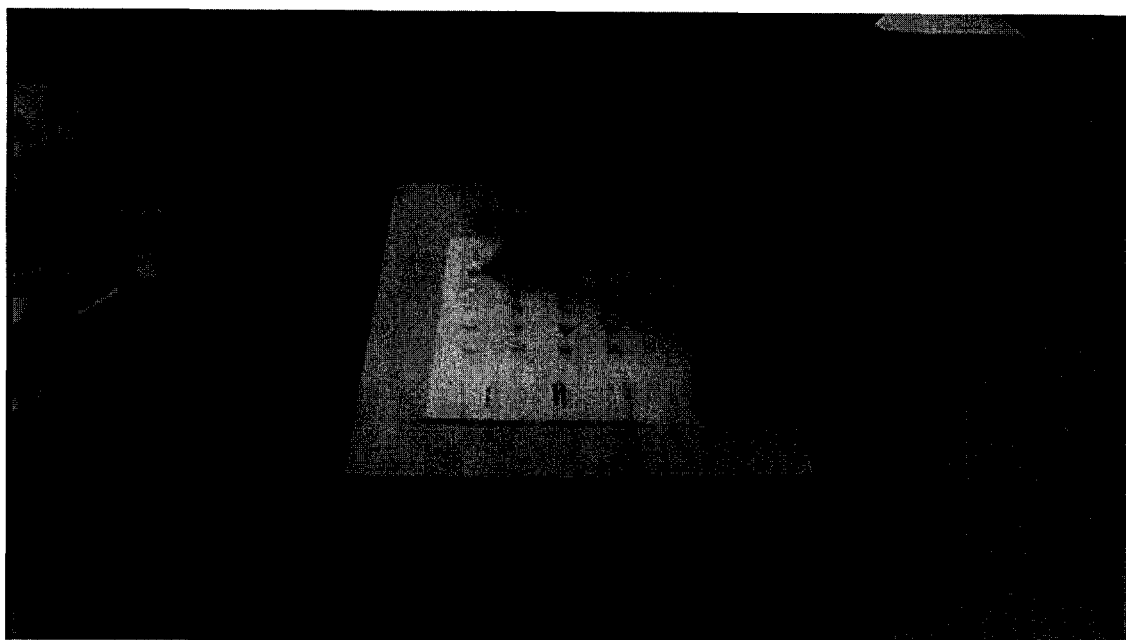


Figure 5.6 A girl helped her classmate by pointing at the letter B when writing the word
Bear
output after “Pizza” has been written, while figure 5.8 shows the output Image of the
word “Cat”.

The output of some words has created a fun atmosphere among the children. For instance, when writing the word “Apple”, a picture of a weird apple with teeth and tongue was displayed. When children saw this image, we have heard a laughter sound and many comments regarding the picture. For example, Anna told the teacher “I don’t think this is real”. Moreover, Peter was laughing and said “This apple is hungry”. Figure 5.9 shows the output image of the word “Apple”. The word ‘Pomme’ shown in the black box is actually the French translation of the word apple.

Not only had the images shown create fun, but also the words that some children have written were really peculiar and humorous. For example, when Tim was asked which word he likes to write, he said “Foot Long”. Once children have heard Tim’s answer, they have laughed and started giving comments.

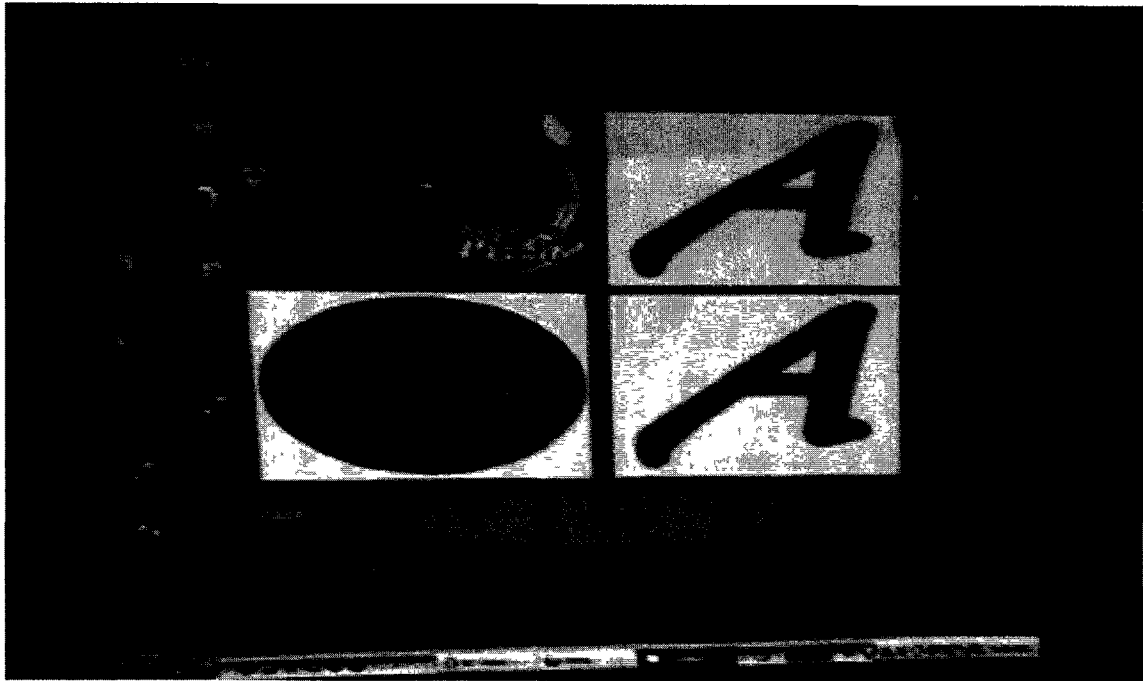


Figure 5.7 The output of the word Pizza

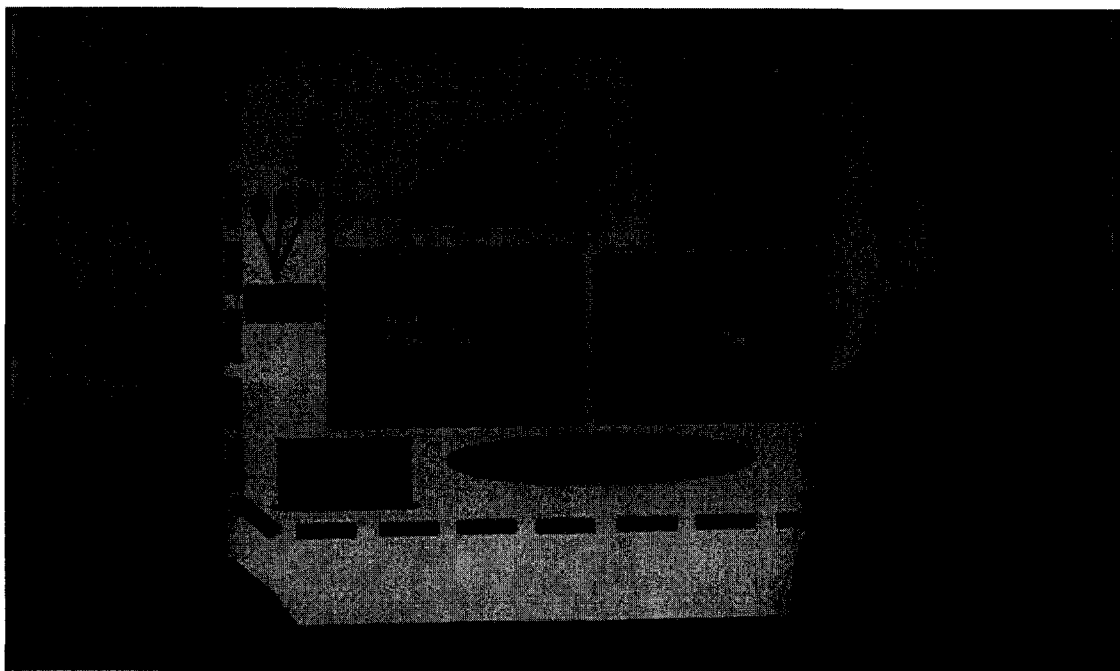


Figure 5.8 The output of the word "Cat"

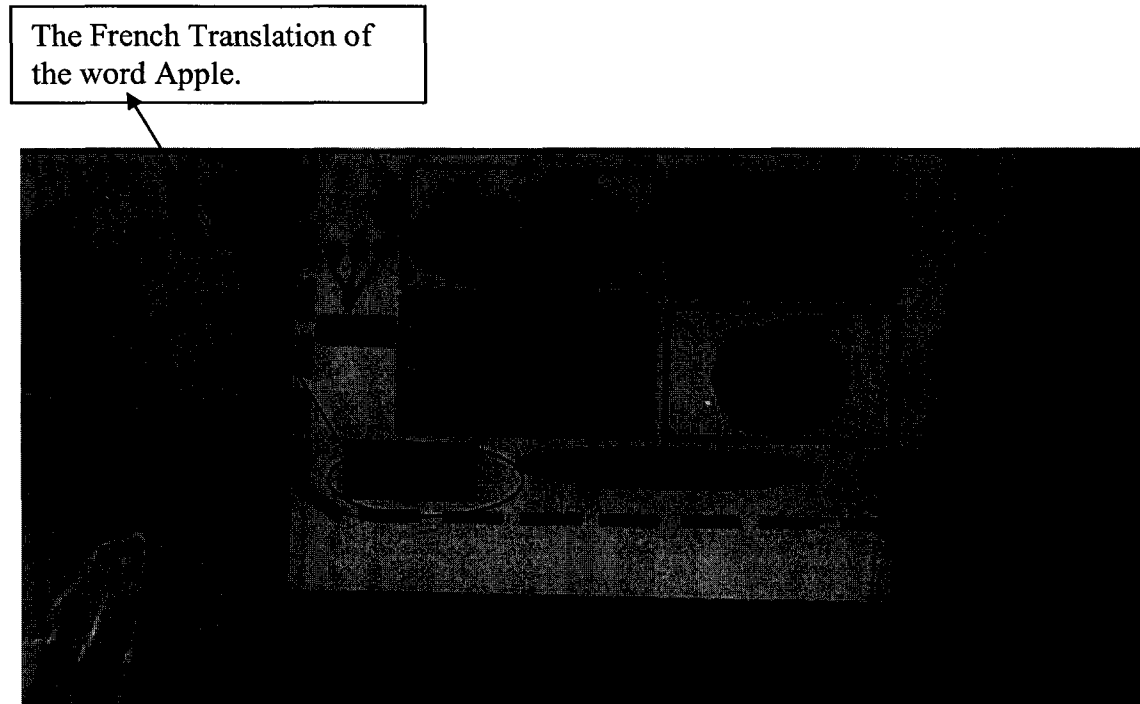


Figure 5.9 The left bottom apple image has caused an atmosphere of laughter among children

What did the children ask? : At the beginning, the teacher was mainly choosing the words to be written by the children since they were a little bit confused about the fact that weird shape stick (Magic Stick) can write letters. Eventually, the confusion was eliminated and they started picking up their own words. For example, one child asked the teacher if it is possible to write the word “Pumpkin”. Another asked if he could write the word “Crocodile”. The output of some pictures also urged the children to pose more questions at the teacher. For instance, when mark wrote “Bear”, the system displayed a white bear and another black. Then he asked her “is there a white bear?” The teacher explained that the white lives in the cold places like Antarctica and the other one lives in the forest. Figure 5.10 shows a child asking the teacher if he could write a word.

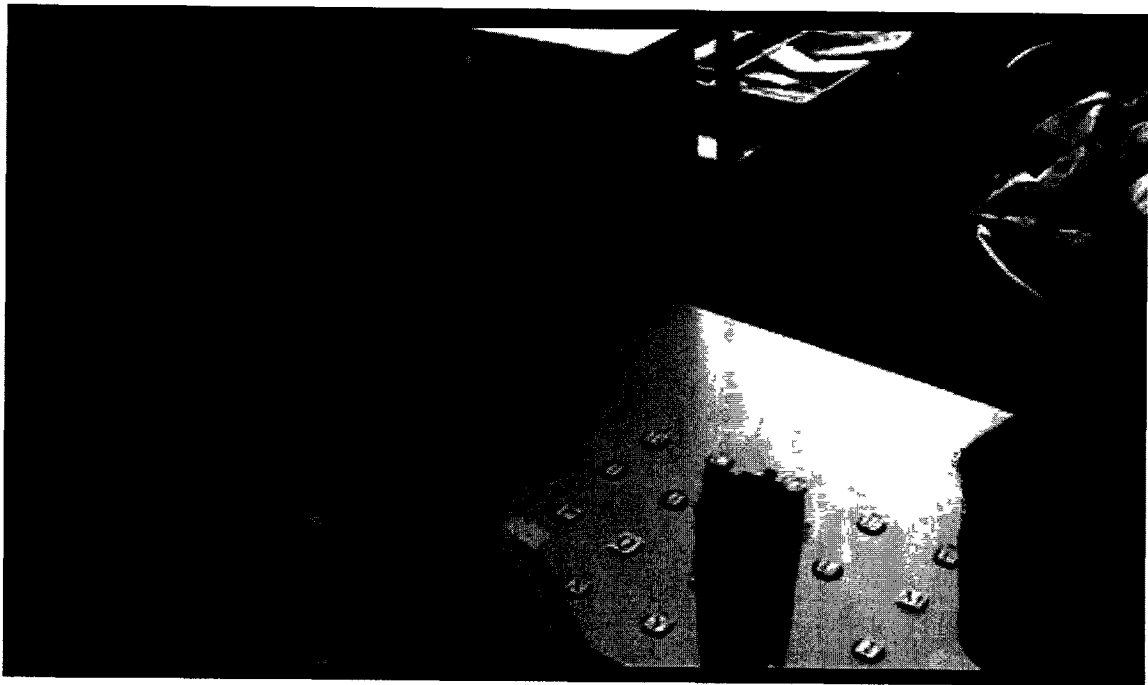


Figure 5.10 A child asking the teacher if he can write “Pumpkin”

Simplicity of tapping letters: Even though children were instructed not to tap very fast and hard on the letters, some of them did not actually allow enough time for ID detection by tapping more than once in a quick and hard manner. This was due to the fact that some children got so excited and they wanted to write quicker. During the 2 experimental sessions, children had tapped a total of about 304 times on the letters. They have succeeded to correctly write a letter from the first tapping attempt almost 246 times, from the second attempt 40 times and from the third try 18 times. Nobody actually required more than 3 tapping attempts in order to write a letter. Figure 5.11 shows the success rate of the number of successful taps made from the first, second and third attempt. The success of correctly tapping a letter from the first attempt was about 81%, while the success rate for the second was almost 13%, and about 6% for the third attempt. This actually would mean that 81% of the times, the children were able to write

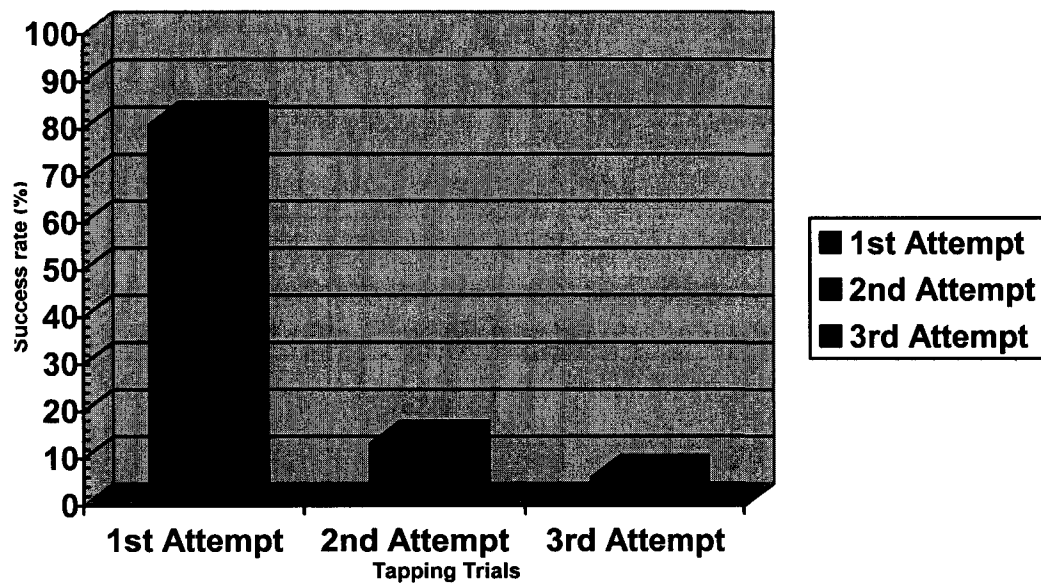


Figure 5.11 Success rate of each of the three Attempts

a letter properly without the need to re-tap on the same letter, which means in our opinion that our Magic Stick performed very well among kindergarteners.

5.2.2 The Object Identification Scenario

- **Experimental Setup**

We have connected a projector for media display, as well as large speaker, to a Laptop where the system's software was running. We tagged some dolls by attaching RFID tags on the top of them. A total number of 20 RFID tagged dolls of animals, birds and fruits were scattered in the classroom.

- **Experimental Procedure**

We conducted qualitative evaluations with the same groups of children of the

previous scenario. We have given 2 Magic Sticks for each group, and we asked the teacher to ask the children to look for different categories and species of animals. The evaluations were conducted in two sessions. Each session took approximately 45 minutes. The sessions were videotaped for further observations.

- **Test Cases**

The teacher have asked each groups children (2 children per group) to look for four different particular objects/entities related to a particular category. For instance, the teacher has asked one group to look for 4 different types of animals that are not a type of a bird. Each group was asked to collaborate in solving the quiz. The objectives of the Object Identification game can be summarized by the following:

- a) Observe how children learned from the Object Identification game about new entities by analyzing their queries.
- b) Observe children's collaboration when solving some short quizzes and learning about new objects and the way they interacted with each other by examining their discussions and dialogues.

- **Results**

Similarly to the Alphabet Learning game, we have based our analysis on the objectives in set in the previous sections for Object Identification game. In this game however, the output was mostly focused on the audio output of the game. Since children were roaming in the classroom and finding objects, it was not always possible for them to look at pictures and hear the names especially from some far corners. For this reason, the teacher was sometimes re-saying what the animated agents were saying. The teacher was indicating whether a choice is wrong or right and guiding the children by explanations and answering their inquiries. In this experiment, we have observed the following:

What did children ask? : The Object Identification game was a good opportunity for the children to pose questions to their teachers. For example, Rachel asked the teacher “what is this grey animal? It seems like a dog”. The teacher responded “this is a wolf, it lives in the forest”. Another girl asked the teacher “is that a squirrel”. And then the teacher responded that that animal is a chipmunk.

Collaboration: Children have assisted each other in finding the objects that the teacher asked them to look for. For instance when Joseph tapped on the duck, his classmate Sara told him “A duck can fly, choose the frog. Not only the children playing in the games have participated, but also children who were watching have contributed in solving of the games by giving suggestions and leading their classmates to the right objects. The following is a transcript of typical conversation between two children A and B when being asked to find 4 types of animals that do not eat herb. This dialogue illustrates children’s discussions and collaboration when solving the game:

Child A: I think a dog eats grass.

Child B: I don’t think so, but let’s choose it.

Child A: Oh he doesn’t! (After teacher told the answer)

Child B: I will touch this goat. I have seen once a goat eating grass in the movie.

Child B: Yummy, this is correct. (After teacher told the answer)

Child A: it is my turn now. Can you please look if there is a cow here? (Asking child B)

Child B: I can’t find it. Think of another animal.

Child A: I don’t have anything on my mind.

Teacher: why don’t you try the rabbit?

Child A: Ok.

Child B: A rabbit like Bugs Bunny.

Child A: it is correct! (After teacher told the answer)

As can be seen from this dialogue, children have collaborated in finding what everyone thinks is the best answer. The comments of the teacher were crucial and provided a fun method of learning for the children.

5.3 Summary

This chapter presented the evaluation of our proposed edutainment system by evaluating both the Magic Stick and voice-based techniques. The evaluation sessions were done in a home and kindergartener school settings among children of ages 2-5. The results pointed out the strength and weakness points of the voice-based and Magic Stick interaction techniques associated with the system. These results may be taken into consideration when further developing the prototype.

Chapter 6

Conclusion and Future Work

This thesis presented our ambient edutainment system that was designed to educate the young children in a simple and entertaining manner. Interaction among children and the system is done through either a voice-based technique or a special user interface called the Magic Stick.

The system allows children to spell, write, and tap on objects and entities, and get in response an appropriate set of media such as images, texts, audios and videos. The system incorporates three different games: the Book game, the Learning Alphabet game, and the Object identification game that promotes learning through fun. We have evaluated our system in both home and kindergartener school environments with children of the ages 2-5.

The Book game was evaluated in a home setting using both the voice-based and Magic Stick techniques. When using the voice-based technique, we had faced difficulties in urging the children to speak some words. In addition, the system didn't recognize all the terms uttered by the children, and as the results have shown, the detection error was about 48%. However, children have performed pretty well when interacting with the

Magic Stick technique. They have grasped easily the concept of tapping with the Magic Stick on the images of the book. We have observed that children were excited during the experiments and have really enjoyed playing. The results obtained when playing the Book game through the Magic Stick have shown that children have successfully tapped and got appropriate media every time the Magic Stick has successfully detected a tag ID. We have realized that some children had to tap more than once on the picture in order to get a response. The rate of success of detecting the tag ID from the first tapping attempt made by the children was 67%, 13.6% from the second, 11.6% from the third, and 7.8% from the fourth attempt.

Now the Learning alphabet and Object Identification games were both evaluated among kindergarteners in a school setting. Children have used the Magic Stick to interact with the system. On one hand, the observations we have made during the Alphabet Learning game has showed that children had really enjoyed interacting with the system. This was obvious from their writings, conversations, discussions, and the questions they have posed to their teachers. On the other hand, the Object identification game has created an atmosphere of collaboration among the children. They have collaborated in finding the correct objects and entities that their teachers have asked them to look for. They also had the chance to learn about new entities through the questions that they have posed to their teachers.

We have noticed during the Object Identification experiment that some children could not hear the names of the objects they were tapping because they were standing sometimes in the corners of the classroom far from the computer's speaker. This issue could be taken into consideration by integrating a small wireless speaker with the Magic Stick so that children at any position can clearly hear the names of the objects they touch. In addition, as was suggested by some teachers, we intend to associate a Math game for teaching children the basic arithmetic operations.

Bibliography

- [1] A. Johnson, M. Roussos, J. Leigh, C. Vasilakis, C. Barnes, and T. Moher, "The NICE project: learning together in a virtual world", *Proceedings IEEE Virtual reality annual international symposium*, 1998, pp.176-183.
- [2] M. Cavazza, R. Aylett, K. Dautenhahn, C. Fencott, and F. Charles, "Interactive Storytelling in Virtual Environments: Building the "Holodeck"", *Proceedings of VSMM*, 2000, pp. 678-687.
- [3] Z. Zhou, A. D. Cheok, and J. Pan, Y. Li, "Magic Story Cube: an interactive tangible interface for storytelling", *Proceedings of the 2004 ACM SIGCHI Singapore*, 2004, pp. 364 – 365.
- [4] F. Garzotto, and M. Forfori, "FaTe2: Storytelling edutainment experiences in 2D and 3D collaborative spaces", *Proceedings of the 2006 conference on interaction design and children*, Tampere, Finland, 2006, pp. 113-116.
- [5] J. P. Hourcade, B. Bederson, A. Druin and G. Taxen, "KidPad: collaborative storytelling for children" *In CHI '02 Extended Abstracts on Human Factors in Computing Systems*, Minneapolis, Minnesota, USA, April 2002, pp. 500-501.
- [6] K. Ryokai, S. Marti, and H. Ishii,"I/O brush: drawing with everyday objects as ink", *Proceedings of the SIGCHI Vienna*, Austria, 2004, pp. 303 – 310.
- [7] H. Raffle, C.Vaucelle, R. Wang, and H. Ishii ,” Jabberstamp: embedding sound and voice in traditional drawings”, *Proceedings of the 6th international conference on Interaction design and children*, Aalborg, Denmark, 2007, pp. 137-144.

- [8] G. Chipman, A. Druin, D. Beer, J. A. Fails, M. L. Guha, S. Simms ,” A case study of tangible flags: a collaborative technology to enhance field trips”, *Proceedings of the 2006 conference on Interaction design and children*, Tampere, Finland, 2006, pp. 1 – 8.
- [9] J. C. Read, “Jabberwocky: children's digital ink story writing from nonsense to sense”, *Proceedings of the 3rd international conference on Digital Interactive Media in Entertainment and Arts*, Athens, Greece, 2008, pp. 85-90.
- [10] A. W. B. Furtado, T. P. Falcão, A.S. Gomes, C. E. M. Rodrigues, and R. Sonnino, “e-du box: educational multimedia with tangible-enhanced interaction”, *Proceedings of the 7th ACM conference on Designing interactive systems*, Cape Town, South Africa, 2008, pp. 139-146.
- [11] Your baby Can Read; <http://www.yourbabycanread.com/>
- [12] Baby’s Learning Laptop; <http://www.vtechkids.com/product.cfm?productID=469>
- [13] LeapSter2; <http://www.leapfrog.com/gaming/leapster2/>
- [14] K. Subrahmanyam, R. E. Kraut, P. M. Greenfield and E. F. Gross, “The Impact of Home Computer Use on Children’s Activities and Development”, *The Future of Children*, Vol. 10, No. 2, 2000.
- [15] J. Meyrowitz, “No sense of Place: The Impact of Electronic Media on Social Behavior”, *Oxford University Press*, 1986.
- [16] FUNBRAIN; <http://www.funbrain.com>
- [17] The Kidz Page; <http://www.thekidzpage.com/>
- [18] D. Xu, E. Mazzone, S. MacFarlane, “Informant Design with Children-Designing Children’s Tangible Technology”, *International Workshop Re-Thinking Technology in Museums*, Limerick, Ireland, 2005.
- [19] Z. Okan, “Edutainment: is Learning at risk?”, *British Journal of Education Technology*, Volume 34, Issue 3, 2007, pp. 255-264.
- [20] Q. Wang, W. Tan, and B. Song, “Research and Design of Edutainment, Information Technologies and Applications in Education”, *ISITAE 07*, 2007, pp. 502-505.

- [21] A. Druin, and C. Solomon, "Designing multimedia environments for children", *New York: John Wiley & Sons*, 1996.
- [22] G. Friedland, W. Hurst, and L. knipping, "Educational Multimedia Systems: The past, the present, and a glimpse into the Future", *Proceedings of the international workshop on Educational multimedia and multimedia education*, ACM, Augsburg, Bavaria, Germany, pp. 1-4, 2007.
- [23] N.S. Said, "An engaging multimodal design model. Interaction design and children", *Proceedings of the 2004 conference on Interaction design and children: building a community*, Maryland, 2004, pp. 169 – 172.
- [24] H.S .Horace, and K. Belton, "Smart Ambience Games for Children with Learning Difficulties", *Technologies for E-Learning and Digital Entertainment*, Springer, Berlin, Heidelberg, 2006, pp. 484-493.
- [25] F. Colace, M. De Santo, and N. Gagliardi, "Multimedia Learning in Advanced Computer-based Contexts: 'Discovering Trier'", *3rd International Conference on Information and Communication Technologies: From Theory to Applications*, IEEE Xplore, Damascus, pp. 1- 4, 2008.
- [26] M. Rehm, E. Andr'e, B. Conradi, S. Hammer, M. Iversen, E. L'osch, T. Pajonk, and K. Stamm, "Location- Based interaction with children for edutainment", *Perception and Interactive Technologies*, Springer Berlin, Heidelberg, 2006, pp. 197-200.
- [27] Z. Muda and I.S. Basiron, "Multimedia Adventure Game as Edutainment Application", *International Conference on Computational Intelligence for Modelling, Control and Automation*, IEEE, Vienna, 2005, pp. 497 -500.
- [28] K. Ryokai and J. Cassell, "StoryMat: a play space for collaborative storytelling", *Conference on human factors in computing systems. CHI'99 extended abstracts on human factors in computing systems*, Pittsburg, Pennsylvania, 1999, pp. 272-273.
- [29] M. Bllinghurst, H. kato, and I. Poupyrev, "MagicBook: Transitioning between reality and virtuality", *Conference on Human Factors in Computing Systems*, Seattle, Washington, 2001, pp. 25-26.

- [30] W. Dietz, Jr. Gortmaker, "Do we fatten our children at the television set? Obesity and television viewing in childhood and adolescents", *Pediatrics*, pp. 807-812.
- [31] M. Griffiths, "Friendship and social development in children and adolescents: The impact of electronic technology", *Educational and Child Psychology*, Vol 14, 1997, pp. 25-37.
- [32] J. Colwell, C. Gradyand, S. Rhaiti, "Computer games, self esteem, and gratification of needs in adolescents", *Journal of Community and Applied Social Psychology*, Vol 5, 1995, pp. 195-206.
- [33] S.L. Calvert, V. J. Rideout, J. L. Woolard, R. F. Barr and G. A. Strouse," *Age, Ethnicity and Socioeconomic Patterns in Early Computer Use: A National Survey, American Behavioral Scientist*", 2005.
- [34] A. Karime, M. A. Hossain, W. Gueaieb and A. El Saddik, " Magic Stick: A Tangible Interface for the Edutainment of the Young Children", *IEEE International Conference on Multimedia and Expo (ICME2009)*, June 28- July 3, 2009, NY, USA.
- [35] A. Karime, M. Anwar Hossain, A. El Saddik and W. Gueaieb, "A Multimedia-Driven Ambient Edutainment System for the Young Children ", *ACM MM 2nd Workshop on Story Representation, Mechanism and Context (SRMC08)*, Vancouver, Canada, 2008, pp. 57-64.
- [36] J. Reicher, D. Reicher and M. Reicher, "Use of Radio Frequency Identification (RFID) Tags in Bedside Monitoring of Endotracheal Tube", *Journal of Clinical Monitoring and Computing*, Springer, Netherlands, Vol 21, 2007, pp. 155-158
- [37] N. Erasala and D. C. Yen, "Bluetooth technology: a strategic analysis of its role in global 3G wireless communication era", *Computer Standards and interfaces*, Vol 24, issue 3, 2002, pp. 193-206.
- [38] M. Pressley, "Imagery and Children's Learning: Putting the Picture in Developmental Perspective", *Review of educational Research*, Vol 47, 1977, pp. 585-622.
- [39] W. A. Nelson, "Efforts to improve computer-based instruction: the role of knowledge representation and knowledge construction in hypermedia systems", *Computer in the schools*, Vol 10, 1995, pp. 371- 399.

- [40] F.W. Grover, "Inductance Calculations Working Formulas and Tables", *Dover Publications Inc*, 1946.
- [41] A.M Rahman, A. El Saddik, "Traffic Architecture Driven Organization and Visualization of Learning Objects Metadata in Virtual Environment," *3rd Annual E-learning Conference on Intelligent Interactive Learning Object Repositories (I2LOR)*, Montreal, Quebec, 2006.
- [42] Wordnet- a lexical database for the English Language;
<http://wordnet.princeton.edu/>
- [43] D. D. Shade, and J. A, "In mother's lap: The effect of microcomputers on mother teaching behavior and young children's classification skills", *Paper presented at the biennial meeting of the Society for Research in Child Development*, Toronto, Canada, 1985.
- [44] C.B. Johnston, "Interactive storybook software". *Early Child Development and Care*, Vol 132, pp. 33-44.
- [45] L. Deng, and X. Huang, "Challenges in adopting speech recognition", *Multimodal interfaces that flex, adapt and persist*, Communications of the ACM, Vol 47, 2004, pp. 69-75.
- [46] Windows Speech Recognition; <http://www.microsoft.com/windows/windows-vista/features/speech-recognition.aspx>
- [47] Google Ajax Language API; <http://code.google.com/apis/ajaxlanguage/>
- [48] Microsoft Agent; <http://www.microsoft.com/msagent/>
- [49] M.K. Shields and R.E. Behrman, "Children and Computer Technology: Analysis and Recommendations", *The Future of Children, Children and Computer technology*, Princeton University, Vol 10, No.2, 2000.
- [50] C# Custom Shape Region:
http://www.vcskicks.com/custom_shape_form_region.html.