

# Katha-Mala: A Voice Output Communication Aid for the Children with Severe Speech and Multiple Disorders (SSMI)

Arijit Mukhopadhyay  
Saptarshi Biswas

Pratik Worah  
Ramasish Das

Susmit Biswas  
Anupam Basu

*Dept. of Computer Science and Engineering, IIT Kharagpur*

## Abstract

*Recent developments in embedded systems technology have opened up a vast area of research and development- the development of portable and affordable assistive devices tuned to specific applications for the physically challenged populace. In this paper we present Kathamala, a novel voice recording and playback system that has been designed to help people with Severe Speech and Multiple Disorders (SSMI). The system is fully portable and has been designed to cater to the Indian community. The biggest advantage of the system is its inherent user-friendliness achieved with the incorporation of appropriate interfacing circuitry. Aided with its inherent flexibility to accept inputs from a variety of sources Kathamala is a complete system on board solution for children who deserve special attention.*

## 1. Introduction

Assistive technology [1, 2] deals with the development of technology hardware and software that enables physically challenged people to overcome the hindrances in efficient communication. Many of the modern assistive devices are computer based [3, 4]. However, for better usage, it is required that the systems be on low cost portable platforms. The affordability issue is all the more important for the people in developing countries, since most of the portable systems are costly with respect to these target users. The portability issue therefore calls for indigenous development of low cost embedded assistive systems.

The necessity for the development of a system like Kathamala came into our mind from the needs expressed to us by the spastic children at the Indian Institute of Cerebral Palsy, Kolkata. About 60% of the students affected with cerebral palsy are speech impaired. Gifted

with all other facilities they feel severely impaired to express.

Kathamala is a Voice Output Communication Aid (VOCA). Such Voice Output Communication Aids are available in the market abroad but they are extremely costly and unaffordable by Indian standards. As an illustrative example, TecTalk [5] and BullShooter [6] cost around \$400. Such devices use specialized DSPs for playing 16 pre-recorded messages.

Hence, it was imperative for us to come up with a design that was indigenous and at the same time very much affordable. Compared to the aforementioned devices our device not only come with a voice recording and playback facility but also has an extremely user-friendly GUI.

As is the case with the design of any embedded system we have presented a complete Hardware Software Co-design. The core of the system is a simple 8031 microcontroller. It has been interfaced with a low cost ASIC, that serves the purpose of an audio-recording/on board memory. Analog preamplifiers have been added at the input stage to make the system compatible with mike-level voltages. The output of the system is applied to an audio amplifier for the audio signals to be played on a standard 8 Ohm speaker. Appropriate Hardware-Software partitioning has been incorporated in the design of the interfacing of the switches and the ASIC and the various I/O components.

The paper is organized as follows.

Section 2 presents the overview of the hardware platform. The next section describes the user view and the directions for usage of the system. Section 4 presents the modes of usage of the system and Section 5 illustrates the two most important features of the system the indigenous Graphical User Interface (GUI) and the access mechanism. The next section presents the comparison of the design with other available ones

followed by a brief overview of the circuit design. Section 7 concludes the paper.

## 2. Hardware Platform

At the heart of the design is the ISD4004 ChipCorder® [7] which provides 3 volt single-chip record/playback solutions for 8-16 minute messaging applications which are ideal for a portable system like the one this paper describes. The ISD4004 is ideal for use in embedded environment. Here the ISD4004 is interfaced with an Intel 8031® [8] microcontroller. Recordings are stored in on-chip non-volatile memory cells thus facilitating zero power message storage. This system uses an ISD4004-8M chip which has maximum message duration of 8 minutes. It uses a sample rate of 8 KHz and a filter pass band of 3.4 KHz. A 4 wire (SCLK, MISO, MOSI, and SS) Serial Peripheral Interface is provided. The pre-amplifier and the output audio amplifier section has been designed to provide good quality voice reproduction at low costs.

## 3. User View

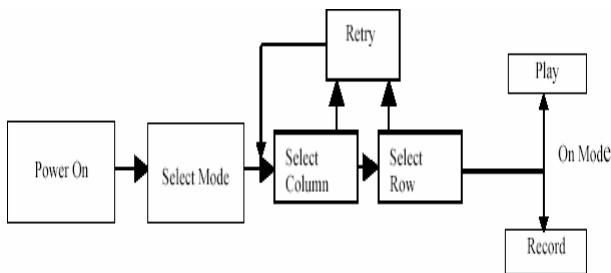


Figure 1

The user view of the system is shown in Figure 1. After the system power is turned the user, through special user friendly switches can select a specific memory location from which he can playback/record depending on the Mode selected.

## 4. Modes of Usage

The system provides high quality playback solutions for 8-16 minute messaging applications. The system has been programmed in such a manner that it can be driven in 5 different modes which can store messages with durations of 1,2,4,8 and 16 minutes. For this purpose the

system is equipped with extremely easy to use multifunctional switches. The switches are basically of 2 types access **switches** and management **switches**. Management **switches have been designed for operation** by the parents, teachers etc. which perform the function of Power-on, Reset, selecting the mode of operation of the device such as play/record or the message durations. **Access** switches have been designed for use by the children by virtue of which they can select the particular location that they want to play from. Another novel feature of the system has been the auto-scan mechanism which allows the user to browse through the locations before locking on one to play from.

## 5. Graphical User Interface (GUI)

The system has been equipped with a user-friendly GUI a first among voice output systems.



Figure 2

An icon will be selected from the icon gallery using the access mechanism. The controller will select the voice or generate the speech corresponding to the icon selected

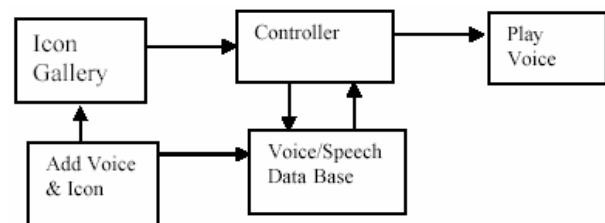


Figure 3

and will feed it to the Play-Voice module as shown in Figure 3. The Add Voice and Icon Module can be used to update the voice database and the icon gallery.

## 6. Access Mechanism

A snapshot of the fully integrated system is shown below in Fig. 4. Children, with motor disabilities that do not allow voluntary control over the fingers or hands, fail to use the direct keyboards. Thus the software and hardware for the icon gallery has been modified to make it a scanning display with suitable timing delay. The elements of the icon gallery can be lit up in a sequence. And the auto-scan facility the icons scan over at the predetermined pace and sequence.

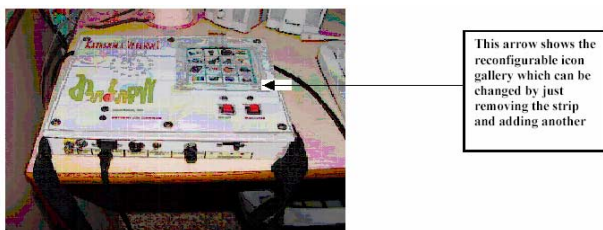


Figure 4

The selection of the desired icon is done through a single touch switch, shown in Fig. 5. This has been implemented with appropriate hardware interfacing and interrupt service routines in the software code.

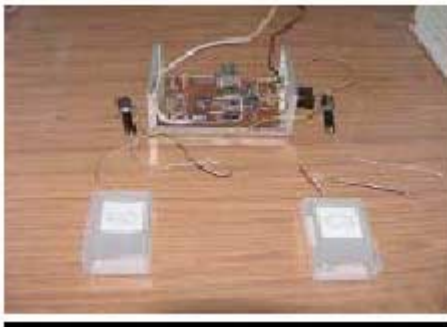


Figure 5

This device in its form has been field-tested at the IICP. The biggest advantage has been the reduction in the on board memory requirements and the removal of any sort of redundancy in the system i.e. a set of icons which are applicable in the market place may be totally redundant in the school. Another striking facet of this system has solved the accessibility problem of children suffering from motor-neuron impairments.

## 7. Design of the System

Figure 6 shows the circuit diagram of the system excluding the preamplifier and the audio amplifier which are shown separately. The biggest advantage of using the ISD4004 ChipCorder® is that it can be directly interfaced with the 8031 microcontroller by using the 4 wire (SCLK, MISO, MOSI, and SS') Serial Peripheral Interface. Another indigenous characteristic is the bank of LEDs. This arrangement is like the table data structure wherein each LED is a reference to a particular memory location in the ISD. This LED array provides a very comfortable interface for the target users.

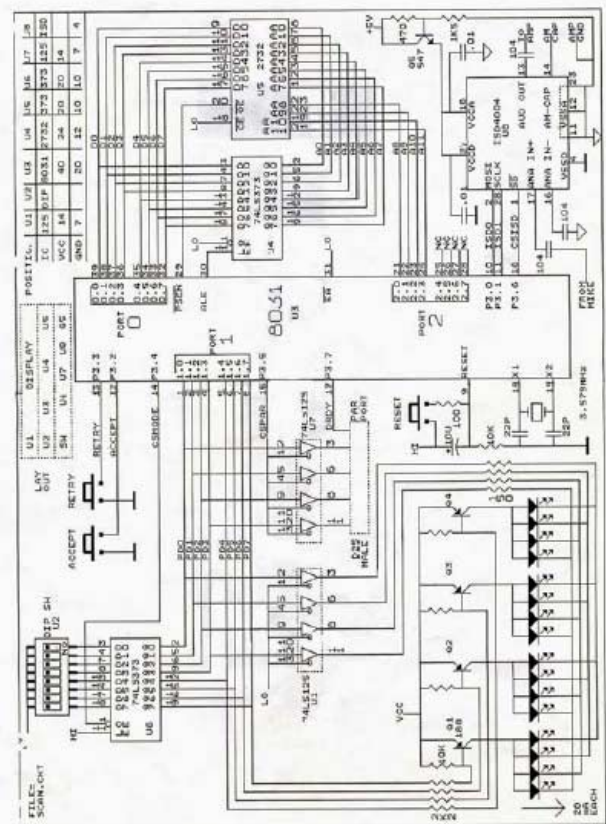


Figure 6

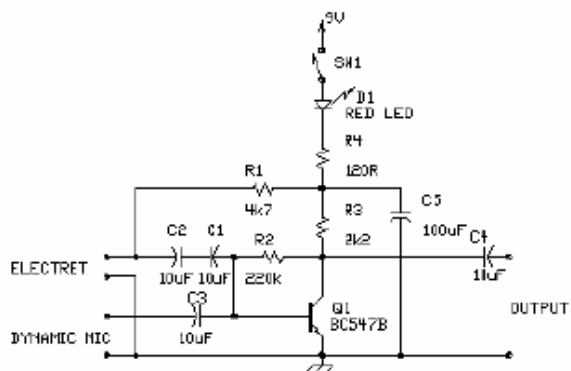


Figure 7

The audio preamplifier is shown in Figure and the audio amplifier is shown in Figure 8.

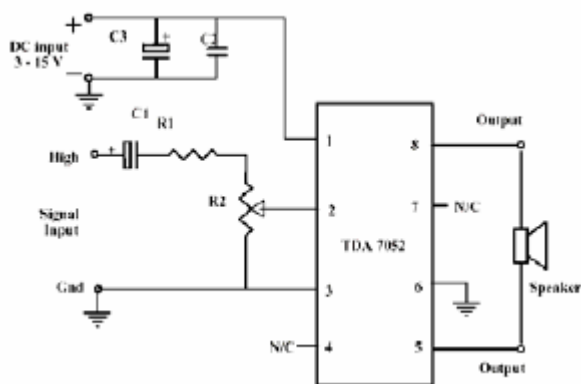


Figure 8

## 8. Comparison

Several such voice output devices are commercially available. Typical examples of such “Light Tech” devices are Speakeasy by Ablenet [voice memory of 4 minutes for 12 switches, costing around \$700.00], Big Mack by Ablenet [single message of 20 second duration, costing \$100.00], Parrot by Zygo [voice memory of 32 seconds, for 16 messages, costing \$1200.00]. A very good reference to the products is available in [1]. Obviously, such costly devices are beyond the affordability of the users in the developing countries including India. In response to this need, researchers at IIT Kharagpur and Microware Technologies, India have developed KathaMala ( garland of voices), which can store messages up to 8 minutes spread over 16 switches. For 16 messages, each message can be of 30 seconds

duration. The system is reconfigurable to enhance the duration of the messages at the cost of the number of messages. The biggest advantage of this system is that it is a low cost solution for the cerebrally challenged students.

## 9. Conclusion

KathaMala-The Garland of Words- has been implemented using innovative design methodologies and appropriate hardware/software partitioning of the system which has helped reduce the cost of the device to Indian affordability. But a point that needs to be mentioned is that to break the cost barrier by other such devices at no point has there been a compromise with the quality of the system. All through the design it has the target user group for this system has been kept in mind. Thus we have tried to stick to the specifications that were fixed from the suggestions put forward by IICP and at the same the novel architecture of the system has aided in keeping the price of the system to less than Rs. 4000/-.

## 10. References

- [1]. Proceedings of the *ACM SIGCAPH Conference on Assistive Technologies*, 2000.
- [2]. Vanderheyden P B. and Pennington C.A, An augmentative communication interface based on conversational schemata, in *Assistive Technology and Artificial Intelligence, Applications in Robotics, User Interfaces, and Natural Language Processing*, (Mittal V.O et al eds.), *LNAI 1458*, Springer Verlag, 1998.
- [3]. Newell Morris A, Booth L, and Ricketts I, Syntax PAL: A writing aid for language-impaired users, in *ISAAC-92*. [Also in *Augmentative and Alternative Communication*, Vol.8, 1992].
- [4]. Pennington C.A. and McCoy K.F, Providing Intelligent Language Feedback for Augmentative Communication Users, in *Assistive Technology and Artificial Intelligence, Applications in Robotics, User Interfaces, and Natural Language Processing*, (Mittal V.O et al eds.), *LNAI 1458*, Springer Verlag, 1998
- [5]. Rumble Gilian and Larcher Janet, *AAC Device Review*, Vocation Pub. 1998
- [6]. [www.ramseyelectronics.com](http://www.ramseyelectronics.com)
- [7]. [www.winbond-usa.com](http://www.winbond-usa.com)
- [8]. [www.ececs.uc.edu](http://www.ececs.uc.edu)