A Reading Tool for the Illiterate

Sandeep Jain
www.sjain.net
sj@sjain.net

ABSTRACT
A handheld scanner can be adapted to help illiterate people teach themselves how to read, and to read printed text out to them. To make this useful, we need to make the hardware cheap enough, create effective optical character recognition and text-to-speech software in the user's language, and come up with a business case that would give a manufacturer the incentive to market this device.

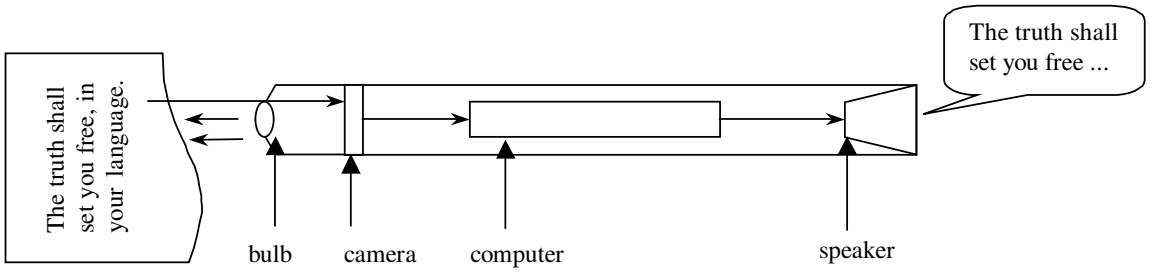
Keywords
Literacy, reading tool, handheld scanner, learning

INTRODUCTION
This paper describes a recent commercial product that can be turned into a tool to help an illiterate person read and learn from printed text. That product is the handheld scanner, a computer scanner packaged in a pen like shape. It can speak out the words on a printed page, as one draws it across them. People who cannot read for medical reasons – because of difficulties in seeing or in recognizing alphabets – have this device speak the page out to them.

Illiterate people can do the same. The product sells for about $200 in the US. If we can make and sell the product for $20, or Rs. 1000, in India, it gives an interested illiterate person a pathway to knowledge. If we can reduce the cost much beyond that, illiterate people may try it out of casual interest. It could be distributed widely. Reducing the cost is the central, but solvable problem.

A LEARNING TOOL
An illiterate person is one who can speak a language, but cannot read it. The person can have this device speak out a page of printed text. It is also a good tool for teaching the person how to read, that is, how to associate the marks on a printed page with the sounds that contain meaning. One would do this by running the device across a single alphabet or a single word and hearing it spoken out. This way, a person, with the device in hand, can teach oneself the sounds of the alphabets, and how these come together to form the sounds that one associates with words. The value of this is that the user will eventually be able to read on his or her own, without the need of a reading aid. The problem here is that the same alphabet can sound different in different words (for example, the c's in 'a piece of cake?'). But in most Indian languages, like Hindi, which are based on what is called the 'Brahmi' script, such differences in the way the same alphabet can sound are minor.

HOW IT WORKS
The handheld scanner has a light at the front that illuminates the text under it. A very basic digital camera behind the bulb takes in the reflected light to form snapshots of the text. These snapshots are fed into a miniature computer behind the digital camera, which translates written text into corresponding speech. It feeds this speech into a little speaker, which speaks out the sentence. The arrangement would look like this:



How it works

The camera and computer shown above work according to the same principles as the expensive digital cameras and computers one buys. We can think of making them very cheap, because they only need to be capable of capturing and processing text. The computer shown above runs software that does the following jobs:

1. Convert the *photograph* of a sentence fed in by the camera into a *sentence of alphabets* that the computer can work with. There is a subtle difference between these two. A computer stores, say, the letter 's' inside itself as a unique sequence of 0's and 1's, in a total of 16 such binary digits. Sixteen zeroes and ones can be flipped in 65,536 different ways (a sequence of three binary digits can take 8 different values). We can match every letter in the alphabet of every reasonably well-known language in the world to one of these 65,536 values. That lets a computer uniquely identify it. However the physical appearance of the letter 's' can take as many forms as different printing presses choose to give it. The job of this first piece of the software is to translate the 's' shape in the photograph, into its unique 16 binary digit code. This is called optical character recognition. This needs to be done for the language of interest to us, say Hindi.

2. Given a sentence in a certain language, translate that sentence from text (that is, a sequence of alphabets) into another sequence of computer words that encode how the sentence would sound if it were spoken out loud. Text-to-speech software exists for many languages. A text-to-speech program was recently published for Hindi by a group of undergraduate students at Punjab University [1].

3. Run these computer 'sounds' through a device called a digital to analog converter, which converts the 1's and 0's into electrical voltages of the kind that the speaker unit expects, and send these voltages to the speaker. The speaker unit then speaks the sentence out to the user.

REDUCING THE COST

The IRIS Pen II Scanner[2] from I.R.I.S. does all of the above and costs $189. (Price is as quoted on the Amazon website, in October 2002). This and other commercial scanners perform a host of other functions, such as keeping a dictionary in memory, editing scanned text in the scanner itself, connecting to other computers, and so on. We need to define our functionality – the minimum is what I have described above, and assemble the most basic hardware that will do the job.

The microcontroller and memory that make up the computer are available for $5 to $10 each. I have not come across an imaging chip (what is described as the camera above) in that price range. It needs to be found or invented. The light bulb and speaker are inexpensive, and the software is for us to write.

ACKNOWLEDGMENTS

I would like to thank Aaron Szymanski and Nitin Sawhney for their encouragement and help, as the work proceeds on trying to turn this into a reality.

REFERENCES

1. http://deepti.nourl.org
   Deepti – a Hindi Speaking Chat Robot
2. http://www.irisusa.com
   I.R.I.S. Optical Character Recognition Products
3. http://www.simputer.org
   The Simputer: A People's Computer
4. http://www.cpen.com
   Home Page of the C Pen Handheld Scanner